IN-12852

Annual Status Report

submitted to

NASA
Goddard Space Flight Center
Greenbelt, Maryland

Error Control Techniques for Satellite and Space Communications
Grant Number NAG5-557
June 1, 1985-May 31, 1987

Principal Investigator:

Daniel J. Costello, Jr.
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

N7315423

July, 1986

Summary of Results

In this report, we analyze the performance of several high rate convolutional inner code/RS outer code concatenated coding systems for possible use in NASA's high speed digital satellite communication systems. Three types of inner codes are considered:

1) Rate $(n-1)/n$, $n>3$, punctured convolutional codes which produce a total effective information rate between 0.5 and 1 bits per unit bandwidth. Punctured codes are decoded using the same decoding trellis as NASA's standard rate 1/2 inner codes. However some parity bits are deleted, resulting in a higher effective rate. This causes some trellis branches to contain two bits (one information and one parity), while other branches contain only one information bit, and may cause problems in properly synchronizing the decoder.
2) High rate majority-logic decodable convolutional codes. These codes also produce an effective information rate between 0.5 and 1 bits per unit bandwidth, and the decoders are simple and capable of high speed operation. However, performance falls far short of what can be achieved with Viterbi decoding.
3) Bandwidth efficient trellis codes using MPSK modulation which achieve effective information rates greater than 1 bit per unit bandwidth. These codes can make use of soft decision Viterbi decoding to achieve significant coding gains on bandlimited channels.

Two decoding schemes are considered. Scheme 1 is a concatenated coding system without side information. The received sequence is first decoded by the inner decoder using the conventional Viterbi algorithm. The outputs of the inner decoder are then grouped into bytes, deinterleaved, and decoded by

an outer errors-only RS decoder. Our results indicate that coding gains from 5 to 8 dB can be achieved at decoded BER's of 10^{-6} to 10^{-9} with little or no bandwidth expansion. For example, with a 64-state punctured inner code and an overall code rate of 0.75, coding gains of 7.75 dB @ 10^{-9} and 5.05 dB @ 10^{-6} are achievable. With a 4-state bandwidth efficient inner code, 8-PSK modulation, and an overall code rate of 1.0, coding gains of 7.15 dB @ 10^{-9}, and 5.05 dB @ 10^{-6} are achievable. This 4-state decoder is capable of high speed operation.

Scheme 2 is a concatenated coding system with side information. The received sequence is first decoded by the inner decoder with a modified Viterbi algorithm. In this modified algorithm, "path metric" comparisons are used to estimate the entire information sequence and "branch metric" comparisons are used to provide side information on the estimated sequence. This is done by erasing the estimated bits that are probably in error. The outputs of the inner decoder are then grouped into bytes, deinterleaved, and decoded by an outer errors-and-erasures RS decoder. If the erasure-correcting capability of the outer code is exceeded, the block is erased. Our results indicate that extremely low undetected bit error rates (BER's) can be achieved with only moderate erasure probabilities at signal-to-noise ratios near 5dB and with little or no bandwidth expansion. For example, undetected BER's as low as 10^{-14}, with an erasure rate no more than 10^{-3}, can be achieved with a 64-state punctured inner code with overall code rate 0.765 at an E_b/N_0 of 5.35 dB or with a 16-state bandwidth efficient inner code, 8-PSK modulation, and an overall code rate 0.9 at an E_b/N_0 of 5.45 dB. The 16-state decoder would be a factor of 4 faster than the 64-state decoder.

HIGH RATE CONCATENATED CODING SYSTEMS

Robert H. Deng

Daniel J. Costello, Jr.

Dept. of Elec. & Comp. Engr.

Univ. of Notre Dame

Notre Dame, IN 46556

Abstract

In this paper, we consider high rate concatenated coding systems with trellis inner codes and Reed-Solomon (RS) outer codes for application in satellite communication systems. Two types of inner codes are studied.

1) High rate punctured binary convolutional codes which result in overall effective information rates between 1/2 and 1 bit per channel use;

2) Bandwidth efficient signal space trellis codes which can achieve overall effective information rates greater than 1 bit per channel use.

Channel capacity calculations with and without side information are carried out for the concatenated coding system. It is shown that, for moderate values of constraint length, optimal performance is obtained by choosing the inner code rate near inner channel cutoff rate.

Two concatenated coding schemes are investigated. In Scheme I, the inner code is decoded with the Viterbi algorithm and the outer RS code performs error-correction only (decoding without side information). In scheme II, the inner code is decoded with a modified Viterbi algorithm which produces reliability information along with the decoded output. In this algorithm, path metrics are used to estimate the entire information sequence, while branch metrics are used to provide the reliability information on the decoded sequence. This information is used to erase unreliable bits in the decoded

output. An errors-and-erasures RS decoder is then used for the outer code. These two schemes are proposed for use on NASA satellite channels. Our results indicate that high system reliability can be achieved with little or no bandwidth expansion.

I. Introduction

Concatenated coding has long been used as a practical means of achieving reliable communication over satellite channels [1,2,3]. One such system consists of Viterbi decoding of an (n,1,m) convolutional inner code concatenated with an outer Reed-Solomon (RS) code. The overall effective information transmission rate of this system is less than 1/n, $n \geq 2$, bits per channel use [2,3]. This system achieves increased power efficiency for decreased bandwidth efficiency. With the ever-increasing demand for satellite communication services, both available channel bandwidth and transmitter power must be conserved. Thus, the search for bandwidth and power-efficient modulation/coding systems has recently become a very active research area.

In this paper, we present a trellis code/RS code concatenated coding system for use in high speed satellite communication systems. Two types of trellis codes are considered:

1) High rate punctured binary convolutional codes with code rates $(n-1)/n$, $n \geq 3$ [4,5].

2) Bandwidth efficient signal space trellis codes with effective code rates of 1 or greater [6,7,8,9].

The advantage of using punctured codes is that decoding is simplified, especially if the decoder must be capable of decoding rate 1/n and rate k/n codes. With rate $(n-1)/n$, $n \geq 3$, punctured codes as inner codes, an overall effective information rate greater than 1/2 but less 1 bit per channel use can be achieved. The goal of using bandwidth efficient signal space trellis codes as inner codes is to achieve overall effective information rates around 1 bit per channel use. As we will see in the following sections, such a system can provide a rather large coding gain at bit error rates (BER's) of 10^{-6} ~ 10^{-9} with no bandwidth expansion.

Fig. 1 shows the encoding-decoding block diagram of the concatenated coding system. Encoding is performed in two stages. An information sequence of Kb bits is divided into K bytes of b bits each, and each b-bit byte is regarded as a symbol in $GF(2^b)$. These K bytes are used as the input to the RS encoder. The output of this encoder is an N-byte codeword which is byte-interleaved and then serially encoded by the trellis encoder. Decoding is accomplished in the reverse order. The output of the maximum-liklihood (Viterbi) trellis decoder contains bursty errors. The main purpose of the inner code is to shape the distribution of the errors on the inner channel, rather than to correct errors, particularly when the inner code rate is around the cutoff rate of the inner channel. When an RS outer code is used, this shaping compresses the random errors on the inner channel into symbol errors corresponding to the symbol size used by the outer code [10]. The use of bandwidth efficient signal space trellis codes as inner codes has one important feature: because of the bandwidth efficient property of the code, it compensates for the bandwidth expansion introduced by the outer RS code, so that the overall system suffers no bandwidth expansion.

The organization of the paper is as follows. In section II, the outer channel is modeled as a block interference (BI) channel [11]. We then calculate the channel capacity of the outer channel both with and without side information. This calculation shows that:

1) Concatenated code performance (in terms of outer channel capacity) is optimized at a certain inner code rate, denoted by R_1^{opt}. The value of R_1^{opt} depends on the inner code constraint length, it increases as the constraint length increases, and it approaches the inner channel capacity as the constraint length goes to infinity. For moderate constraint lengths, R_1^{opt} is close to the inner channel cutoff rate.

2) For a given inner code constraint length, there exists a limit on the inner code rate, denoted by R_1^{max}, beyond which coding becomes useless (the outer channel capacity goes to zero).

3) As the inner code rate increases from R_1^{opt} to R_1^{max}, system performance (outer channel capacity) degrades rapidly. This helps explain the steep BER curves of concatenated codes.

In sections III - V we study two concatenated coding schemes. Scheme I, a Viterbi decoded trellis inner code/RS outer code system, where the Viterbi decoder outputs (without side information) are provided to the outer decoder, is studied in section III. A modified Viterbi decoding algorithm for trellis codes with output reliability information is presented in section IV. Scheme II, a concatenated coding system with side information available to the outer decoder, is then studied in section V. In the modified algorithm, path metrics are used to estimate the entire information sequence, and branch metrics are used to provide reliability information on the decoded sequence. This is done by erasing a certain number of bits on the survivor of a comparison between the two most likely paths whose metrics are within some preset threshold of each other. Finally, in section VI, we summarize our results and draw some conclusions which are useful in the design of concatenated coding systems.

II. Channel Capacity Calculations

We assume the inner channel to be a discrete memoryless channel (DMC). After decoding the inner code, however, the outer channel (the channel seen by the outer decoder) is no longer a memoryless channel. It has been transformed into a nonuniform or time varying channel by the inner decoder. In this section, we will consider the outer channel formed by the byte interleaver, inner encoder, inner channel, inner decoder, and byte deinterleaver, as shown in Fig. 1. We assume that an (n,k,m) convolutional code is used as the inner code. We first show that this outer channel can be model as a BI channel [11]. It has been argued that channel capacity is the best measure of system performance, whereas cutoff rate is an inverse measure of decoding delay [11]. This observation is further supported by noting (from (3.7) and (4.17) in [11]) that channel capacity is additive while cutoff rate is not for BI channels with side information. Therefore, to assess concatenated code performance, we will calculate the channel capacity of the outer channel.

It has been shown that [12], for any DMC, there exists an (n,k,m) time-varying convolutional code for which the probability of an event error of length j + m branches with maximum-likelihood decoding is upper bounded by

$$P_j(\underline{q},\rho) < \left[(2^k-1)2^{(j-1)k}\right]^{\rho} 2^{-(m+j)kE_0(\underline{q},\rho)/R_1},$$

$$\text{for } 0 \le R_1 < C_1,\ 0 \le \rho \le 1,\ j = 1,2,\ldots, \qquad (1)$$

where $R_1 = k/n$ is the inner code rate, C_1 is the inner channel capacity, and

$$E_0(\underline{q},\rho) = -\log_2 \sum_y \left[\sum_x q(x)p(y|x)^{1/(1+\rho)}\right]^{1+\rho}. \qquad (2)$$

Since in a j + m-branch error event the last m k-bit information tuples must be correct for the incorrect path to remerge with the correct path, the error burst is confined to the first $m(j) \triangleq jk$ bits of the error event. When a code with symbols over $GF(2^b), b \geq 2$, is used as the outer code, a j+m-branch error event can cause up to $t(j) \triangleq \lceil m(j)/b \rceil + 1$ symbol errors, where $\lceil x \rceil$ denotes the integer greater than or equal to x. Hence, the symbol error probability in the outer channel can be upper bounded by

$$P_s(\underline{q},\rho) < \sum_{j=1}^{\infty} t(j)\, P_j(\underline{q},\rho)$$

$$< \sum_{j=1}^{J-1} t(j)\, ((2^k-1)2^{k(j-1)})^{\rho} 2^{-(M+m(j))E_0(\underline{q},\rho)/R_1}$$

$$+ \sum_{j=J}^{\infty} \left(\frac{kj}{b} + 2 \right) \left((2^k-1)2^{k(j-1)} \right)^{\rho} 2^{-(M+kj)E_0(\underline{q},\rho)/R_1},$$

$$\text{for } 0 \leq R_1 < C_1,\ 0 \leq \rho \leq 1, \tag{3}$$

where J, $1<J<\infty$, is arbitrary and $\nu \triangleq mk$ is the encoding constraint length*. Let the second term on the right hand side of (3) be denoted by $Q(\underline{q},\rho)$. Then

$$Q(\underline{q},\rho) = (2^k-1)^{\rho} 2^{-(\nu+k)E_0(\underline{q},\rho)/R_1} \left\{ \sum_{j=J}^{\infty} \frac{kj}{b} 2^{k\rho(j-1)} 2^{-k(j-1)E_0(\underline{q},\rho)/R_1} \right.$$

$$\left. + \sum_{j=J}^{\infty} 2 \cdot 2^{k\rho(j-1)} 2^{-k(j-1)E_0(\underline{q},\rho)/R_1} \right\}$$

$$= (2^k-1)^{\rho} 2^{-(\nu+k)E_0(\underline{q},\rho)/R_1} \left\{ \frac{k}{b} \sum_{j=J}^{\infty} jz^{j-1} + 2 \sum_{j=J}^{\infty} z^{j-1} \right\}, \tag{4}$$

where $z \triangleq 2^{k\rho} 2^{-kE_0(\underline{q},\rho)/R_1}$. To obtain a closed form expression for

*For simplicity, we assume that each encoder shift register has full length m.

$Q(\underline{q},\rho)$, note that

$$\sum_{j=J}^{\infty} z^{j-1} = \frac{z^{J-1}}{1-z} \quad , \qquad |z| < 1,$$

and by taking the derivative of both sides we obtain

$$\sum_{j=J}^{\infty} jz^{j-1} = \frac{d}{dz}\left(\sum_{j=J}^{\infty} z^{j}\right) = \frac{d}{dz}\left(\frac{z^{J}}{1-z}\right) = \frac{z^{J}}{(1-z)}\,2 + \frac{Jz^{J-1}}{1-z} \quad , \qquad |z| < 1.$$

After some manipulation, (4) can now be rewritten as

$$Q(\underline{q},\rho) = (2^{k}-1)^{\rho}\left\{\left(\frac{kJ}{b} + 2\right)\frac{wz-1}{1-z} + \frac{k}{b}\,\frac{w}{(1-z)^{2}}\right\}, \quad 0 \le \rho\, E_0(\underline{q},\rho)/R_1 \le 1, \tag{5}$$

where $w \triangleq 2^{k\rho J}\, 2^{-(\nu + kJ + k)E_0(\underline{q},\rho)/R_1}$.

Substituting (5) into (3) and minimizing $P_s(\underline{q},\rho)$ over $\underline{q}$ and ρ, we obtain the following closed form expression for the symbol error probability:

$$P_s < \min_{\substack{\underline{q} \\ 0\le\rho\le 1}} \left\{\sum_{j=1}^{J-1} t(j)\,((2^{k}-1)2^{k(j-1)})^{\rho} 2^{-(\nu+m(j))E_0(\underline{q},\rho)/R_1} + Q(\underline{q},\rho)\right\},$$

$$0 \le R_1 < C_1. \tag{6}$$

Due to the interleaving, symbol errors in the outer channel are statistically independent, and hence the outer channel is equivalent to a BI channel with two memoryless BSC's, as shown in Fig. 2. Channel Δ_0 is a noise-

less BSC, and $\Delta_{1/2}$, assuming the worst case, is a useless BSC with crossover probability 1/2. A sequence of b consecutive bits is sent over Δ_0 with probability $Pr\{\Delta_0\} = 1 - P_s$ and over $\Delta_{1/2}$ with probability $Pr\{\Delta_{1/2}\} = P_s$.

The outer channel capacity with and without side information is given by [11]

$$\overline{C}_2 = 1 - Ps \tag{7}$$

and

$$C_2 = (1-p) - \frac{1}{b} \{H(p) + p \log_2 (1-2^{-b})\}, \tag{8}$$

respectively, where $p \overset{\Delta}{=} P_s(1-2^{-b})$ and H(p) is the binary entropy function. Then the overall channel capacity of the concatenated coding system is

$$\overline{C} = \overline{C}_2 \cdot R_1 \quad \text{(with side information)} \tag{9}$$

and

$$C = C_2 \cdot R_1. \quad \text{(without side information)} \tag{10}$$

Note that the capacity without side information approaches the capacity with side information as b increases, but the rate of convergence is slow (o(1/b) convergence). Also, from (6) we see that P_s is a decreasing function of b; therefore, both $\overline{C}$ and C increase as b increases.

To obtain some numerical results, assume the inner channel is a BSC with crossover probability $\varepsilon = 0.1$. Fig. 3 shows the overall channel capacity of the concatenated coding system, normalized by the inner channel capacity C_1, for k=3 and b=8, with ν as a parameter. For fixed ν, there is an optimum value of R_1/C_1 that maximizes the overall capacity. We denote the corresponding inner code rate by $\overline{R}_1^{opt}$ (R_1^{opt}). Note that the optimum inner code rate does not occur at the value of R_1 which presents the best (least noisy)

channel to the outer decoder. Instead, there is a local maximum for each value of ν that drifts towards the higher rates as ν increases. This is reasonable since the system performance should improve for increasing contraint length. For moderate values of ν, $\overline{R}_1^{opt}(R_1^{opt})$ is close to the inner channel cutoff rate. In addition, for a fixed ν, there is a limit on R_1, denoted by $\overline{R}_1^{max}$ (R_1^{max}), beyond which the overall capacity goes to zero (for $\nu=3$, $\overline{R}_1^{max}=0.60C_1$ and $R_1^{max}=0.58C_1$). Therefore, the inner code rate is constrained between 0 and $\overline{R}_1^{max}(R_1^{max})$. In fact, in a concatenated coding system, the inner code rate often lies between $\overline{R}_1^{opt}(R_1opt)$ and $\overline{R}_1^{max}$ (R_1^{max}). The rapid increase in overall capacity as R_1 goes from $\overline{R}_1^{max}(R_1^{max})$ to $\overline{R}_1^{opt}(R_1^{opt})$ corresponds to the sharp decrease in BER with increasing signal power (and hence C_1) in well designed concatenated coding systems.

III. Coding Scheme I - Without Side Information

A concatenated coding system is depicted in Fig. 1. Decoding is accomplished in the reverse order of encoding. The inner decoder uses the Viterbi algorithm. The outputs of the inner decoder (without side information) are grouped into b-bit symbols. Because the lengths of the bursts of output errors made by the Viterbi decoder are widely distributed, b-bit symbol deinterleaving is used so that errors in the individual RS-symbols of one block (N symbols) are independent; otherwise, a very long block code would be required to operate the system efficiently.

Let the minimum distance of the (N,K) RS code be d_2. When an N-symbol block is received by the outer decoder, it performs errors - only decoding. That is, if the N-symbol block contains $t_2 = \left\lfloor \frac{d_2-1}{2} \right\rfloor$ or fewer errors, the errors

are corrected; otherwise, decoding fails. t_2 is called the error-correcting-capability of the (N,K) RS code [13]. The decoded bit error rate at the output of outer decoder can be closely approximated by

$$P_b \cong \frac{d_2}{2N} \sum_{i=t_2+1}^{\infty} \binom{N}{i} (P_s)^i \ (1-P_s)^{N-i}, \tag{11}$$

where P_s is the symbol error probability into the outer decoder.

In this section, we give several examples of the Scheme I concatenated coding system. For each example, we use RS codes with symbols over $GF(2^8)$ and N=255. The symbol error probability P_s is obtained by computer simulation on an AWGN channel unless otherwise noted, and (11) is used to estimate the final decoded BER.

A. Examples with inner punctured codes

Example 3.1:

The best punctured rate R_1 = 3/4, 5/6, and 7/8 codes that are formed from the basic (2,1,6) convolutional code [5] are used as inner codes. Decoding is by the Viterbi algorithm with 3-bit quantization. The decoded BER P_b vs. E_b/N_0 is shown in Figs. 4.1 - 4.3 for 3 different RS code minimum distances. Fig. 4.4 shows the E_b/N_0 required to achieve decoded BERs of 10^{-6} and 10^{-9} as a function of the overall code rate R. Note that, for the given decoded BERs and a given R, lower inner code rates gives better performance. From Fig. 4.4, we see that the best performance is obtained at $R \cong 0.65$ for R_1 = 3/4, and at $R \cong 0.73$ for R_1 = 5/6 and 7/8. These curves indicate the existence of an optimum outer code rate, i.e., the outer code rate which minimizes the required E_b/N_0 to achieve a given decoded BER for a fixed inner code rate.

Fig. 4.5 shows system performance for the case $R_1 = 3/4$, $R = 0.656$, and $d_2 = 33$ vs. R_1/R_{01}. As R_1/R_{01} goes to 1 (the channel becomes noisier), system performance degrades rapidly. This corresponds to the rapid decrease in overall capacity with increasing R_1 illustrated in Fig. 3.

B. Examples with inner bandwidth efficient trellis codes

The goal of using bandwidth efficient trellis codes as inner codes is to achieve a moderate coding gain with no bandwidth expansion. Currently, QPSK is the prevalent modulation scheme in use for digital satellite communications. To improve power efficiency, error-correcting codes have been employed. Because extra bits must be added to the information sequence to perform the error correction, the modulator must operate at a higher rate, thus requiring a larger bandwidth. On the other hand, if bandwidth is limited and the modulator must operate at the same rate as in the uncoded case, the information transmission rate is lowered. To improve power efficiency and maintain a high rate of information transmission without bandwidth expansion, coded modulation, or bandwidth efficient coding, in which trellis codes are combined with an expanded set of modulation signals (typically 8-PSK or 16-PSK), provides an exciting alternative.

In our concatenated coding system, bandwidth efficient inner trellis codes have two functions:

1) To compensate for the bandwidth expansion introduced by the outer RS code;

2) To compress the random errors on the inner channel into symbol errors which can be corrected by byte-error-correcting codes, such as RS codes.

This concatenated coding scheme performs well on channels which are both power and bandwidth limited, as shown in the following examples.

Example 3.2

Periodically time-varying trellis codes (PTVTC) of rate $R_1 = (2P + i)/3P$, $P > 2$, $1 < i < P$, are used as inner codes with 8-PSK modulation [7]. These codes are especially designed for high data rate channels with low decoding complexity. This PTVTC/8-PSK system has an effective information rate $R_{eff}^{(1)} = \frac{3}{2} R_1 = \frac{(2P + i)}{2P} > 1$ bits/channel use, so it is possible to achieve an overall effective information rate R_{eff} equal to 1 bit/channel use or greater. Moreover, because of the periodic property of the code, its trellis structure is the same as the trellis structure of a lower-rate code, and as a result the complexity of the decoder is reduced significantly [7].

The PTVTC's are decoded by the Viterbi algorithm assuming no demodulator output quantization. Figs. 5.1 - 5.3 show the decoded BER P_b vs. E_b/N_0 for concatenated codes with 4-state inner codes and 3 different RS outer codes; Figs. 5.4 - 5.6 give similar results for 16-state inner codes. In Figs. 5.7 - 5.8 we show the required E_b/N_0 to achieve decoded BERs 10^{-6} and 10^{-9} for 4-state and 16-state inner codes, respectively.

Note that at high effective code rates, e.g., $R_{eff} = 1$ bit/channel use, we still get large coding gains. For example, with a 16-state, $R_1 = \frac{7}{9}$ PTVTC, the coding gain equals 3.92 dB at $P_b = 10^{-6}$ and 6.18 dB at $P_b = 10^{-9}$. Even for a 4-state, $R_1 = \frac{7}{9}$ PTVTC, coding gains of 3.68 dB and 5.98 dB at $P_b = 10^{-6}$ and $P_b = 10^{-9}$, respectively, can be obtained. Only a small increment in coding gain is achieved as the number of trellis states increases. As the number of states goes from 4 to 16, only about 0.2 dB more gain is obtained. This appears to be a characteristic of coded modulation in general [6]. We will see this more clearly when we derive the asymptotic coding gain.

Example 3.3

Ungerboeck's 16-state, rate $R_1 = 2/3$ convolutional code with 8-PSK modulation is used as the inner code. This code has an effective information rate $R_{eff}^{(1)} = 1$ bit/channel use. Therefore, the overall effective information rate, R_{eff}, of the concatenated coding system is less than 1. Again, the inner code is decoded by the Viterbi algorithm without demodulator output quantization. The final decoded bit error probability P_b is shown in Fig. 6.1 for 2 different RS outer codes. Fig. 6.2 shows the required E_b/N_0 to achieve $P_b = 10^{-6}$ and 10^{-9}. At $P_b = 10^{-6}$ and 10^{-9}, coding gains of 4.96 dB and 7.25 dB can be obtained with only 12.5% bandwidth expansion.

Example 3.4

In this example, two kinds of trellis codes are used as inner codes. The first kind uses 8-PSK modulation twice per trellis interval (forming a 64-ary set), and this set is then coded with a $R_1 = 5/6$ trellis code [8]. Viewed more generally, code symbols are mapped onto a four-dimensional signal set (8PSK x 8PSK) with 5 information bits for every two 8-PSK signals. Thus the effective information rate is $R_{eff}^{(1)} = \frac{5}{4} = 1.25$ bits/channel use. The second kind uses rate $R_1 = 3/4$ convolutional codes with 16-PSK modulation [9]. The effective information rate is $R_{eff}^{(1)} = \frac{3}{4} = 1.5$ bits/channel use. This code is more bandwidth efficient than the $R_1 = 5/6$, 8-PSK code, but it is less power efficient.

Let d_f, E_s, and A be the normalized minimum Euclidean free distance, the energy per modulation symbol, and the number of nonzero information bits in the set of error events at distance d_f from the correct path, respectively. Then at high signal-to-noise ratios, the bit error probability at the Viterbi

decoder output is approximated by

$$\varepsilon \cong \frac{A}{5} Q\left(\sqrt{\frac{d_f^2 E_s R_2}{2N_0}}\right) = \frac{A}{5} Q\left(\sqrt{\frac{d_f^2 (5/2) E_b R_2}{2N_0}}\right) = \frac{A}{5} Q\left(\sqrt{\frac{d_f^2 1.25 E_b R_2}{N_0}}\right), \qquad (12.1)$$

for $R_1 = 5/6$ and 8-PSK,

and

$$\varepsilon \cong \frac{A}{3} Q\left(\sqrt{\frac{d_f^2 E_s R_2}{2N_0}}\right) = \frac{A}{3} Q\left(\sqrt{\frac{d_f^2 (3E_b) R_2}{2N_0}}\right) = \frac{A}{3} Q\left(\sqrt{\frac{d_f^2 1.5\ E_b R_2}{N_0}}\right) \qquad (12.2)$$

for $R_1 = 3/4$ and 16-PSK,

where R_2 is the outer RS code rate. The RS-symbol (over $GF(2^b)$) error probability at the output of the Viterbi decoder is bounded by

$$P_s \leq b\varepsilon. \qquad (13)$$

The final decoded bit error probability P_b, obtained from (12.1), (12.2), (13), and (11), is shown in Figs. 7.1 - 7.4, where we have chosen the outer RS code rate R_2 such that the overall effective information rate R_{eff} = 1 bit/channel use. Tables 1.1 and 1.2 list the coding gains at $P_b = 10^{-6}$ and 10^{-9} for the $R_1 = 5/6$, 8-PSK and $R_1 = 3/4$, 16-PSK inner codes, respectively. In both cases, the relatively poor performance of the 8-state and 16-state codes compared to the 4-state codes is due to a large number of minimum free distance paths.

Example 3.5

In this example, we use high-rate self orthogonal convolutional codes with majority-logic decoding as inner codes. These codes have the advantage of an extremely simple and fast inner decoder. Assuming the inner channel to be a BSC with crossover probability p, the BER of the inner decoder output for a self-orthogonal (n,k,m) convolutional code is bounded by

$$\varepsilon \leq \frac{1}{k} \sum_{i=t_{ML}+1}^{n_E} \binom{n_E}{i} p^i (1-p)^{n_E-i} \quad , \tag{14}$$

where

$$p = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad , \tag{15}$$

for BPSK or QPSK modulation on an AWGN channel. n_E is the effective constraint length and t_{ML} is the majority-logic error correcting capability of the code [13]. Using (14), (15), (13), and (11), the final decoded BER is shown in Figs. 8.1 and 8.2 for rate $R_1 = 3/4$ and 4/5 self-orthogonal inner convolutional codes, respectively. Three different values of t_{ML} are considered in each case. Although these codes have implementation advantages, their performance is clearly much worse than the punctured codes with Viterbi decoding.

C. Asymptotic concatenated coding gain

Having studied the performance of concatenated coding systems for small and medium values of E_b/N_0, it is interesting to look at how the systems perform at large values of E_b/N_0 - the asymptotic coding gain. This approach also gives us some insight into concatenated coding system design.

We will use BPSK or QPSK as our reference system in the coding gain calculation. When uncoded BPSK or QPSK is used on an AWGN channel with coherent demodulation, the demodulator output BER is given by

$$P_b = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \cong \frac{1}{2} e^{-E_b/N_0} \quad , \text{ for large } \frac{E_b}{N_0} \quad . \tag{16}$$

1) Asymptotic coding gain with punctured inner codes.

Let A be the number of nonzero information bits on all weight d_f paths for an (n,k,m) convolutional code. If BPSK or QPSK modulation is used, the free Euclidean distance d_f normalized by E_s, the symbol energy at the modulator output, is related to the free Hamming distance $d_f(H)$ by

$$d_f^2 = 4d_f(H), \quad \text{for BPSK,} \tag{17.1}$$

$$d_f^2 = 2d_f(H), \quad \text{for QPSK.} \tag{17.2}$$

At high signal-to-noise ratios with either BPSK or QPSK modulation, an asymptotically tight expression for the bit error probability at the output of a Viterbi decoder without demodulator output quantization is*

$$\varepsilon \cong \frac{A}{k} Q\left(\sqrt{\frac{d_f^2 RE_s}{2N_0}}\right) = \frac{A}{k} Q\left(\sqrt{\frac{2d_f(H)RE_b}{N_0}}\right)$$

$$\cong \frac{A}{2k} e^{-\frac{d_f(H)RE_b}{N_0}}, \tag{18}$$

where $R = R_1R_2$, $R_1 = k/n$ is the inner convolutional code rate, and $R_2 = K/N$ is the outer (N,K) RS code rate with symbols over $GF(2^b)$. From (18), (13), and (11), the final decoded bit error probability P_b of the concatenated coding system can be approximated by

$$P_b \cong \frac{d_2}{N} \binom{N}{t_2+1} P_s^{t_2+1} (1-P_s)^{N-t_2-1}$$

$$\cong \frac{d_2}{N} \binom{N}{t_2+1} P_s^{t_2+1} \cong \frac{d_2}{N} \binom{N}{t_2+1} \left(\frac{Ab}{2k}\right)^{t_2+1} e^{-\frac{d_f(H)RE_b(t_2+1)}{N_0}}. \tag{19}$$

* For QPSK modulation, $E_s = 2E_b$ since there are 2 bits transmitted per symbol.

Comparing (16) to (19), we see that for a fixed E_b/N_0, the (negative) exponent with concatenated coding is larger by a factor of $d_f(H)R(t_2+1)$ than the exponent without coding. Since the exponential term dominates the error probability expressions for large E_b/N_0, we define the asymptotic coding gain as

$$\gamma \triangleq 10 \log_{10} d_f(H)R(t_2+1). \quad \text{(without demodulator output quantization)} \tag{20.1}$$

If hard quantized demodulator outputs are used at the input of the Viterbi decoder, it can be shown that

$$\gamma = 10 \log_{10} \frac{d_f(H)R(t_2+1)}{2}. \quad \text{(with demodulator output quantization)} \tag{20.2}$$

For RS codes, $N - K = 2t_2$ and

$$\frac{N-K}{N} = 1-R2 = \frac{2t_2}{N}. \tag{21}$$

Substituting (21) into (20.1) and (20.2), the asymptotic coding gain can be rewritten as

$$\gamma = 10 \log_{10} d_f(H)R_1R_2\left(\frac{N(1-R_2)}{2} + 1\right)$$

$$= 10 \log_{10} d_f(H)R_1 + 10 \log_{10} R_2\left(\frac{N(1-R_2)}{2} + 1\right), \quad \text{(without quantization)} \tag{22.1}$$

and

$$\gamma = 10 \log_{10} \frac{d_2(H)R_1}{2} + 10 \log_{10} R_2\left(\frac{N(1-R_2)}{2} + 1\right), \quad \text{(with quantization)} \tag{22.2}$$

where the first term is due to the inner convolutional code and the second term depends on the RS outer code. This implies that, for very large E_b/N_0, the inner and the outer codes in a concatenated coding system can be designed

independently of each other.

Taking derivatives of γ with respect to R_2, the maximum coding gain is achieved at $R_2 = \frac{1}{2} + \frac{1}{N}$ for a fixed inner code, and

$$\gamma_{max} = 10 \log_{10} d_f(H)R_1 + 10 \log_{10} \frac{1}{2}\left(\frac{N}{4} + 1 + \frac{1}{N}\right) \quad \text{(without quantization)} \tag{23.1}$$

and

$$\gamma_{max} = 10 \log_{10} \frac{d_f(H)R_1}{2} + 10 \log_{10} \frac{1}{2}\left(\frac{N}{4} + 1 + \frac{1}{N}\right) . \quad \text{(with quantization)} \tag{23.2}$$

Fig. 9 shows the asymptotic coding gain without demodulator output quantization, where we use $R_1 = (n-1)/n$, $n \geq 3$, punctured convolutional codes formed from the (2,1,6) convolutional code as inner codes and an RS code with N = 255 and symbols over $GF(2^8)$ as the outer code. Note that, at high overall code rates, higher inner code rates outperform lower inner code rates, which is in contrast to the case when E_b/N_0 is small. The important fact is that system performance is very sensitive when the outer code rate R_2 is high. A small increase in R_2 results in a large system performance degradation. This is also true for small values of E_b/N_0, as shown in Figures 4.1 - 4.3.

2) Asymptotic coding gain with bandwidth efficient inner codes

Bandwidth efficient codes have been proposed and extensively investigated for high data rate transmission over voiceband telephone channels and band-limited satellite channels with power constraints. Although most of the previous studies on coded modulation for satellite channels have been focused on constant carrier envelope signaling formats such as 8-PSK and 16-PSK modulation, applications of coded amplitude modulation to high speed satellite communication have also been investigated. However, since constant envelope

signals mitigate the non-linear effects of TWT amplifiers and have relatively simple high speed modem implementations, coded MPSK modulation is the preferred modulation scheme for satellite channels. Hence, in the following asymptotic coding gain derivation, we assume that the code is formed from a rate $R_1 = k/n$ convolutional code in combination with 2^L-ary PSK modulation, $L > 2$.

Let E_s be the symbol energy at the modulator output, d_f be the Euclidean distance normalized by E_s, and A be the number of nonzero information bits in the set of paths at distance d_f from the correct path. Then

$$E_s = R_2 R_1 L E_b, \tag{24}$$

and the bit error probability at the output of a Viterbi decoder without demodulator output quantization is

$$\varepsilon \leq AQ\left(\sqrt{\frac{d_f^2 E_s}{2N_0}}\right) = AQ\left(\sqrt{\frac{d_f^2 R_2 R_1 L E_b}{2N_0}}\right)$$

$$\leq \frac{A}{2} e^{-\frac{d_f^2 R_2 R_1 L E_b}{4N_0}}, \tag{25}$$

where $R_1 L$ is the average number of information bits in each modulator output symbol. Using (25) and following a similar procedure as in the punctured inner code case, we can show that the asymptotic coding gain is given by

$$\gamma = 10 \log_{10} \frac{d_f^2 L R_1 R_2 (t_2 + 1)}{4}$$

$$= 10 \log_{10} \frac{d_f^2 \, 2R_{eff}^{(1)} R_2 (t_2 + 1)}{4} = 10 \log_{10} \frac{d_f^2 \, R_{eff} (t_2 + 1)}{2}, \tag{26}$$

where

$$R_{eff}^{(1)} = \frac{L}{2} R_1 \text{ bits/channel use} \tag{27}$$

is the effective inner code information rate, and

$$R_{eff} = R_{eff}^{(1)} R_2 \text{ bits/channel use} \tag{28}$$

is the overall effective information rate of the concatenated coding system.

If an RS code of length N is used as the outer code, by substituting (21) into (26) it follows that

$$\gamma = 10 \log_{10} \frac{d_f^2 R_{eff}}{2} \left(\frac{N(1-R_2)}{2} + 1 \right)$$

$$= 10 \log_{10} \frac{d_f^2 R_{eff}^{(1)}}{2} + 10 \log_{10} R_2 \left(\frac{N(1-R_2)}{2} + 1 \right) , \tag{29}$$

and for a fixed inner code, the maximum coding gain is

$$\gamma_{max} = 10 \log_{10} \frac{d_f^2 R_{eff}^{(1)}}{2} + 10 \log_{10} \frac{1}{2} \left(\frac{N}{4} + 1 + \frac{1}{N} \right) , \text{ with } R_2 = \frac{1}{2} + \frac{1}{N} . \tag{30}$$

Note that in (29) the first term denotes the coding gain due to the bandwidth efficient inner code and the second term represents the coding gain contributed by the outer code; therefore, for very large E_b/N_0, the inner and outer codes can be designed independently in a concatenated coding system. Fig. 10 shows the asymptotic concatenated coding gain γ for Example 3.2, Fig. 11 shows γ for Example 3.3, Fig. 12.1 shows γ for Example 3.4 when $R_1 = 5/6$ coded 8 PSK is used as the inner code, and Fig. 12.2 shows γ for Example 3.4 when $R_1 = 3/4$ coded 16PSK is used as the inner code.

Before finishing this section, we can draw a number of conclusions from the above discussion:

1) For a given overall R_{eff}, there exists an $R_{eff}^{(1)}$ which optimizes system performance; or conversely, for a given $R_{eff}^{(1)}$, there exists an optimum value of R_{eff}.

2) System performance is very sensitive to R_2 at very low and very high values of R_2.

3) For small or moderate values of E_b/N_0, and for a fixed R_{eff}, a lower $R_{eff}^{(1)}$ is preferred as long as the outer code rate R_2 is not too high.

4) Because d_f increases slowly as the number of trellis states increases, choosing inner codes with a small number of trellis states increases the data transmission rate (by reducing the number of decoder computations) with only a slight sacrifice in system performance.

IV. A Modified Viterbi Algorithm

In section II, we have shown that the overall channel capacity of a concatenated coding system with side information is an upper bound on the capacity without side information. The difference between the two capacities is significant, especially on noisy channels, when the inner code rate is in the region from R_1^{opt} to R_1^{max}. We also pointed out in section II that in a concatenated coding system $[R_1^{opt}, R_1^{max}]$ is the practically interesting region for the inner code rate R_1. Therefore, it is advantageous if the inner decoder can provide some kind of reliability information (side information) about its estimated output which will aid the outer decoder. In scheme I, discussed in section III, the inner code is decoded by the Viterbi algorithm and the output of the decoder is a sequence of "hard decisioned" binary digits, i.e., no side information is associated with the output. To provide side information, it becomes necessary to modify the conventional Viterbi algorithm so that soft decisioned outputs are made available to the outer decoder.

Several ideas have been proposed to provide some kind of side information with the convolutional decoder outputs. Zeoli [14] proposed a concatenated coding system that employs a long constraint length (m = 31) convolutional code obtained by annexing a tail to a (3,1,7) convolutional code. The

longer code is then decoded by the same Viterbi decoder as the short code, with the exception that the information sequence along the best path to each state is treated as correct and used to "cancel" the effect of the longer tail from the encoded sequence. The tail provides excellent error-detection capability once the decoder starts to make mistakes. But the algorithm is subject to very serious error propagation and the decoder has to be reset frequently.

Another method of providing reliability information with the decoder output is to compute the a posteriori probability of each decoded symbol from the decoder being correct [3]. However, the a posteriori probability must be computed for each branch using a recursive method and real numbers must also be stored. This computation process slows down the decoder, so it is not suitable for high speed systems.

A third alternative is to use the "Viterbi decoding algorithm for convolutional codes with repeat request" [15] to extract reliability information from the inner decoder. When all the path metrics at some level of the trellis are below a predetermined threshold, the received sequence up to that level is erased (instead of being retransmitted). But this approach has two major drawbacks: 1) when a long received sequence is erased, this erasure information cannot be used by the outer decoder because the number of erasures may exceed the erasure correction capability of the outer code; 2) with high probability, most of the symbols in the received sequence can be decoded correctly and hence should not be erased.

In the following, we propose a decoding technique based on the Viterbi algorithm. It erases only the information symbols that are "probably in error", and this erasure procedure does not affect the decoder's selection of the most likely path. Therefore, the decoder is still maximum-likelihood.

Furthermore, the algorithm is very simple to implement and capable of high speed operation.

As the name suggests, an (n,k,m) trellis code is best described in terms of its trellis diagram. In the decoding of trellis codes by the Viterbi algorithm, a first-event error is made at an arbitrary level j if the correct path is eliminated for the first time at level j in favor of the incorrect path. This is illustrated in Fig. 13. The incorrect path must be some path that had previously diverged from the correct path. For the example shown in Fig. 13, let

$$\underline{V}_j = (\ldots, V_{j-3}, V_{j-2}, V_{j-1}, V_j)$$

and

$$\underline{V}'_j = (\ldots, V'_{j-3}, V'_{j-2}, V'_{j-1}, V'_j)$$

represent the correct path and the incorrect path, respectively, where V_i (V'_i) is the symbol on the ith branch of the correct (incorrect) path, and let

$$\underline{E}_j = (V'_{j-3}, V'_{j-2}, V'_{j-1}, V'_j)$$

represent the 4-branch first-event error at level j. Also let

$$\underline{U}_j = (\ldots, U_{j-3}, U_{j-2}, U_{j-1}, U_j)$$

and

$$\underline{U}'_j = (\ldots, U'_{j-3}, U'_{j-2}, U'_{j-1}, U'_j)$$

represent the information sequences associated with paths $\underline{V}_j$ and $\underline{V}'_j$, respectively, where U_i (U'_i) is a binary k-tuple, and let

$$\underline{I}_j = (U'_{j-3}, U'_{j-2}, U'_{j-1}, U'_j)$$

represent the 4-branch information sequence corresponding to the 4-branch first-event error $\underline{E}_j$ at level j. Denote a branch metric by $\lambda(V_i)$ and a path metric by $\lambda(\underline{V}_j)$. From Fig. 13 we observe that $\underline{I}_j$ is unreliable and should be erased. This observation motivates us to propose the following decoding algorithm.

Algorithm 1:

For each state at level j, select the path $\underline{V}'_j = (\ldots, V'_{j-\ell+1}, V'_{j-\ell+2}, \ldots, V'_{j-1}, V'_j)$ (where ℓ is chosen as the expected error event length) that has the largest metric $\lambda(\underline{V}'_j)$ and the path $\underline{V}_j = (\ldots, V_{j-\ell+1}, V_{j-\ell+2}, \ldots, V_{j-1}, V_j)$ that has the next largest metric $\lambda(\underline{V}_j)$. Because $\underline{V}'_j$ has the largest metric, $\underline{V}'_j$ survives at that state. Moreover, if

$$\lambda(\underline{V}'_j) \geq \lambda(\underline{V}_j) + T \quad (T > 0), \tag{31}$$

decoding continues as in the conventional Viterbi algorithm. On the other hand, if

$$\lambda(\underline{V}'_j) < \lambda(\underline{V}_j) + T, \tag{32}$$

then $\underline{I}_j = (U'_{j-\ell+1}, U'_{j-\ell+2}, \ldots, U'_{j-1}, U'_j)$ on the surviving path is erased and decoding continues.

Note that if T = 0, the algorithm reduces to the conventional Viterbi algorithm. As an illustration of the algorithm, in Fig. 13, for some state at level j, suppose that $\lambda(\underline{V}'_j) > \lambda(\underline{V}_j)$, so the incorrect path $\underline{V}'_j$ survives. If

$$\lambda(\underline{V}'_j) < \lambda(\underline{V}_j) + T, \tag{33}$$

then $\underline{I}_j = (U'_{j-3}, U'_{j-2}, U'_{j-1}, U'_j)$ is erased. Note that (33) is equivalent to

$$\lambda(V'_{j-3}) + \lambda(V'_{j-2}) + \lambda(V'_{j-1}) + \lambda(V'_j) < \lambda(V_{j-3}) + \lambda(V_{j-2}) + \lambda(V_{j-1}) + \lambda(V_j) + T. \tag{34}$$

Eq. (34) implies that, if the metrics of the two 4-branch sequences are "close" in terms of the threshold T, the corresponding information sequence $\underline{I}_j$ is not reliable, and therefore a tag should be attached to it which indicates its degree of reliability. The simplest way of doing this is to erase it.

From (34) we see that this erasure decision is made on the most recent ℓ branches for a ℓ-branch error event, and hence it is a "local" estimate. On the other hand, the path estimate is a "global" estimate. From this point of view, we see that Algorithm 1 is constructed based on the following ideas.

1) By letting the path with the largest path metric survive, we are choosing the maximum-likelihood estimate.

2) By performing a "branch comparison" over the most recent ℓ branches, we provide some reliability information on the maximum-likelihood path.

Because the reliability is estimated over only ℓ branches, the reliability estimates are suboptimum. However, in feedback decoding of convolutional codes, decisions are made over only one constraint length. Hence we can say that the reliability estimates have a certain degree of precision. Ideally, the number of erased k-triples, ℓ, should equal the length of the error event. Practically, however, choosing ℓ equal to the average error event length should be sufficient.

Algorithm 1 can be applied to any trellis code. If, however, the trellis is generated by a finite-state machine, the algorithm can be improved. Let m and k denote the memory order of the finite-state machine and the number of input bits to the finite-state machine. Then trellis codes can be divided into two cases of practical interest:

1) If all k input bits are shifted into memory, there can be no parallel transitions in the trellis. Since the last km information bits associated with an error event of length ℓ branches must be correct for the incorrect path to remerge with the correct path, decoding errors are confined to the first ℓ-m branches.

2) If fewer than k bits are shifted into memory and the remaining bits are uncoded, there are parallel transitions between any two connected states in the trellis, and all the information bits associated with an error event may be in error.

In the case where parallel transitions are possible, we must use algorithm 1. However, in decoding trellis codes without parallel transitions, a modified algorithm can be used.

Algorithm 2.

For each state at level j (j = m + 1, m + 2,...), select the path $\underline{V}'_j = (\ldots, V'_{j-\ell-1}, V'_{j-\ell+2},\ldots,V'_j)$ (where ℓ is chosen as the expected error event length) that has the largest metric $\lambda(\underline{V}'_j)$ and the path $\underline{V}_j = (\ldots, V_{j-\ell+1}, V_{j-\ell+2},\ldots,V_j)$ that has the next largest metric $\lambda(\underline{V}_j)$. Because $\underline{V}'_j$ has the largest metric, $\underline{V}'_j$ survives at that state. Then if

$$\lambda(\underline{V}'_j) > \lambda(\underline{V}_j) + T \quad (T > 0), \tag{35}$$

decoding continues as in the conventional Viterbi algorithm. On the other hand, if

$$\lambda(\underline{V}'_j) < \lambda(\underline{V}_j) + T, \tag{36}$$

then the information symbols $U'_{j-\ell+1}, U'_{j-\ell+2},\ldots, U'_{j-m}$ on the surviving path are erased and decoding continues.

In both algorithms, smaller values of T result in lower erasure rates but higher error rates, and larger values of T result in higher erasure rates but lower error rates.

V. Coding Scheme II - With Side Information

The encoding process is the same as in Scheme I. Decoding is done in two steps. First, the inner trellis code is decoded by the modified Viterbi algorithm presented in the last section. The outputs of the inner decoder consist of binary digits as well as erased bits, and they are grouped into b-bit bytes, deinterleaved, and sent to the outer decoder. Bytes which contain any erased bits are considered as erasures by the outer decoder.

Let i be the number of erased symbols in an RS outer codeword of length N. The outer decoder declares an erasure (or raises a flag) for the entire block of N bytes if i is greater than the erasure-correction threshold T_{es}, where

$$T_{es} \leq (d_2 - 1) \tag{37}$$

and d_2 is the minimum distance of the outer (N,K) RS code. If i is less than T_{es}, the outer decoder starts errors-and-erasures decoding on the N-symbol block, and the i symbol erasures along with the symbol errors are corrected based on the outer code. Let t(i) be the error-correction threshold for a given i, where

$$t(i) \leq \lfloor (d_2-i-1)/2 \rfloor . \tag{38}$$

If the syndrome of the N-byte block corresponds to an error pattern of i erasures and t(i) or fewer symbol errors, errors-and-erasures correction is performed. The values of the erased symbols, and the values and the locations of symbol errors, are determined based on a certain algorithm. However, if more than t(i) symbol errors are detected, then the outer decoder again declares an erasure (or raises a flag) for the entire N-symbol block.

Let P_s and P_e be the RS-symbol error probability and RS-symbol erasure probability at the input of the outer decoder, respectively. The probability of correct decoding of a block is given by

$$P_c = \sum_{i=0}^{T_{es}} \binom{N}{i} P_e{}^i \sum_{j=0}^{t(i)} \binom{N-i}{j} P_s{}^j (1 - P_s - P_e)^{N-i-j} , \tag{39}$$

where T_{es} and t(i) are defined in (37) and (38). Let P_{be} and P_{bi} denote the probabilities of a block erasure and an incorrect decoding, respectively. Then

$$P_c + P_{be} + P_{bi} = 1, \tag{40}$$

and

$$P_{be} + P_{bi} = 1 - P_c$$

$$= \sum_{i=0}^{T_{es}} \binom{N}{i} P_e^i \sum_{j=t(i)+1}^{N-i} \binom{N-i}{j} P_s^i (1-P_e-P_s)^{N-i-j}$$

$$+ \sum_{i=T_{es}+1}^{N} \binom{N}{i} P_e^i (1-P_e)^{N-i} \quad . \tag{41}$$

The probability of incorrect decoding of a block is not easy to determine. But it can be upper bounded by [16]

$$P_{bi} \leq \sum_{i=0}^{T_{es}} \binom{N}{i} P_e^i \sum_{j=d_2-i-t(i)}^{N-i} \binom{N-i}{j} P_s^i (1-P_s-P_e)^{N-i-j} \ , \tag{42}$$

and the final decoded BER of the concatenated coding system is approximated by

$$P_b \cong \frac{d_2}{2N} P_{bi}. \tag{43}$$

In the following we consider several examples for Scheme II. The outer code is again an RS code of length N = 255. Inner trellis codes are decoded by the modified Viterbi algorithm in section IV. We use B to denote the number of erased branches in the modified Viterbi algorithm. The RS-symbol error probability P_s and the RS-symbol erasure probability P_e at the input to the outer decoder are determined by computer simulation. In decoding the outer code, thresholds T_{es} and t, which are independent of the number of erased symbols in an N-symbol block, are assumed, and (41) and (42) are modified as

$$P_{be} + P_{bi} = \sum_{i=0}^{T_{es}} \binom{N}{i} P_e^i \sum_{j=\min(t(i),t)+1}^{N-i} \binom{N-i}{j} P_s^i (1-P_e-P_s)^{N-i-j}$$

$$+ \sum_{i=T_{es}+1}^{N} \binom{N}{i} P_e^i (1-P_e)^{N-i} \qquad (44)$$

and

$$P_{bi} \leq \sum_{i=0}^{T_{es}} \binom{N}{i} P_e^i \sum_{j=d_2-i-\min(t(i),t)}^{N-i} \binom{N-i}{j} P_s^i (1-P_s-P_e)^{N-i-j} \quad , \qquad (45)$$

where $T_{es} \leq d_2-1$ and $t \leq \lfloor (d_2-1)/2 \rfloor$. Usually, the probability of a block erasure P_{be} is much larger than the probability of block decoding error P_{bi}. Therefore, $P_{be} + P_{bi}$ is a tight bound on P_{be}. In the examples below we compute the sum of the probability of a block erasure and of a block decoding error, $P_{be} + P_{bi}$, and the final decoded BER P_b.

A. Examples with Punctured Inner Codes

Example 5.1:

The inner codes are the same as in Example 3.1. In decoding the inner code, we use Algorithm 2. BPSK or QPSK modulation over an AWGN channel with three bit uniform quantization is assumed. The bit metric is given by

$$m(I/1) = \log p(I/1), \quad I = 0,1,2,\ldots,7, \qquad (46)$$

when I is received and 1 is transmitted. In the computer simulation, we use nonnegative bit metrics by letting

$$m(I/1) = \log p(I/1) - \log p(0/1), \quad I = 0,1,\ldots 7, \qquad (47.1)$$

and by symmetry,

$$m(I/0) = m((7-I)/1) , \qquad I = 0,1,\ldots,7. \qquad (47.2)$$

The branch erasure threshold T in Algorithm 2 is chosen to be

$$T = T'm(0/0), \tag{48}$$

where T' is a normalized threshold. Figs. 14.1 - 14.3 show the probabilities P_b and $P_{be} + P_{bi}$ as a function of t.

B. Examples with Bandwidth Efficient Inner Codes

In decoding the bandwidth efficient code we use Euclidean distance as the decoding metric. The erasure threshold in the modified Viterbi algorithm is chosen to be

$$T = (T'd_f E_b/N_0)^2, \tag{49}$$

where d_f is the Euclidean free distance of the code normalized by the signal energy at the demodulator output, and T' is a normalized threshold.

Example 5.2:

As in Example 3.2, the PTVTCs of $R_l = (2P + i)/3P$, $P \geq 2$ and $1 \leq i \leq P$, with 8PSK modulation are used as inner codes. Only 16-state codes are considered. Because of the parallel transitions in the trellis, algorithm 1 is used. P_b and $P_{be} + P_{bi}$, with $R_l = 7/9$, 5/6, 8/9, are shown as functions of t in Figs. 15.1, 15.2, and 15.3, respectively.

Example 5.3:

As in Example 3.3, Ungerboeck's $R_l = 2/3$ coded 8PSK with 16 trellis states is used as the inner code. Because there are no parallel transitions in the code trellis, we use algorithm 2 in decoding the inner code. P_b and $P_{be} + P_{bi}$ are shown in Fig. 16 as a function of t.

From the above examples, we see that for a given E_b/N_0, the decoded BER of Scheme II is significantly lower than that of Scheme I. Scheme II provides us with flexibility in the system design. Tradeoffs between the decoded BER

and the probability of block erasure can be obtained by varying the error correction threshold t of the outer code. Smaller values of t always result in lower decoded BER's, but a higher probability of block erasure. In practice, given a specified system performance, the values of t can be chosen accordingly.

VI. Summary and Conclusions

In this paper we considered high rate concatenated coding systems with trellis inner codes and RS outer codes for application to satellite communication systems. Specifically, we considered high rate punctured binary convolutional inner codes, which result in overall effective information rates between 1/2 bit and 1 bit per channel use, and bandwidth efficient signal space inner codes which can achieve overall effective information rates greater than 1 bit per channel use.

The outer channel in the concatenated coding system was modeled as a BI channel, from which we were able to calculate the overall channel capacities with and without side information. These calculations revealed that for finite total encoder memory order, optimum system performance is achieved at an inner code rate R_1^{opt}, and as the inner code rate increases beyond R_1^{opt}, system performance degrades sharply. We also showed that system performance in terms of channel capacity improves when side information is available.

Two types of concatenated coding systems were studied. Scheme I operates without side information, while Scheme II uses side information. The performance of Scheme I was studied by computer simulations, formula calculations, and by asymptotic coding gain derivations. Results indicated that rather large coding gains could be obtained with little or no bandwidth expansion. In studying the performance of Scheme II, we first proposed a

modified Viterbi algorithm for decoding trellis codes. Side information was provided to the outer decoder in the form of symbol erasures. The outer RS code was then decoded by an errors-and-erasures decoder. A significant improvement in decoded BER was obtained for Scheme II. For systems where block erasures are allowed, Scheme II is highly recommended.

References

[1] G. D. Forney, Jr., Concatenated Codes, Cambridge, MA: MIT Press, 1966.

[2] J. P. Odenwalder, "Optimal Decoding of Convolutional Codes," Ph.D. thesis, University of California, Los Angeles, 1970.

[3] L. N. Lee, "Concatenated Coding Systems Employing a Unit-Memory Convolutional Code and a Byte-Oriented Decoding Algorithm," IEEE Trans. Commun., COM-25, pp. 1064-1074, October 1977.

[4] G. C. Clark, J. B. Cain, and J. M. Geist, "Punctured Convolutional Codes of Rate (n-1)/n and Simplified Maximum Likelihood Decoding," IEEE Trans. on Inform. Theory, IT-25, pp. 97-100, January 1979.

[5] Y. Yasuda, K. Kashiki, and Y. Hirata, "High-Rate Punctured Convolutional Codes for Soft Decision Viterbi Decoding," IEEE Trans. on Commun., COM-32 pp. 315-319, March 1984.

[6] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," IEEE Trans. on Inform. Theory, IT-28, pp. 55-67, January 1982.

[7] F. Hemmati and R.J. F. Fang," Low Complexity Coding Methods for High Data Rate Channels", Comsat Laboratories Technical Note, Feb. 7, 1984.

[8] S. G. Wilson, "Rate 5/6 Trellis-Coded 8PSK", submitted to IEEE Trans. on Commun..

[9] S. G. Wilson, H. A. Sleeper, II, P. J. Schottler, and M. T. Lyons, "Rate 3/4 Convolutional Coding of 16-PSK: Code Design and Performance Study," IEEE Trans on Commun., COM-32, December 1984.

[10] M. A. Herro, D. J. Costello, Jr., and L. Hu, "Capacity and Cutoff Rate Calculations for a Concatenated Coding System," submitted to IEEE Trans. Inform. Theory, July 1985.

[11] R. J. McEliece and W. E. Stark, "Channels with Block Interference," IEEE Trans. Inform. Theory, IT-30, pp. 44-53, January 1984.

[12] A. J. Viterbi and J. K. Omura, Principles of Digital Communication and Coding, New York, NY: McGraw-Hill, 1979.

[13] S. Lin and D. J. Costello, Jr., Error Control Coding: Fundamentals and Applications, Englewood Cliffs, NJ: Prentice-Hall, 1983.

[14] G. W. Zeoli, "Coupled Decoding of Block-Convolutional Concatenated Codes," IEEE Trans. on Commun., COM-21, pp. 219-226, March 1973.

[15] H. Yamamoto and K. Itoh, "Viterbi Decoding Algorithm for Convolutional Codes with Repeat Request," IEEE Trans. Inform. Theory, IT-26, pp. 540-546, Sept. 1980.

[16] T. Kasami, S. Lin, and T. Takata, "A Cascaded Coding Scheme for Error Control," submitted to IEEE Trans. Inform. Theory, Jan. 1986.

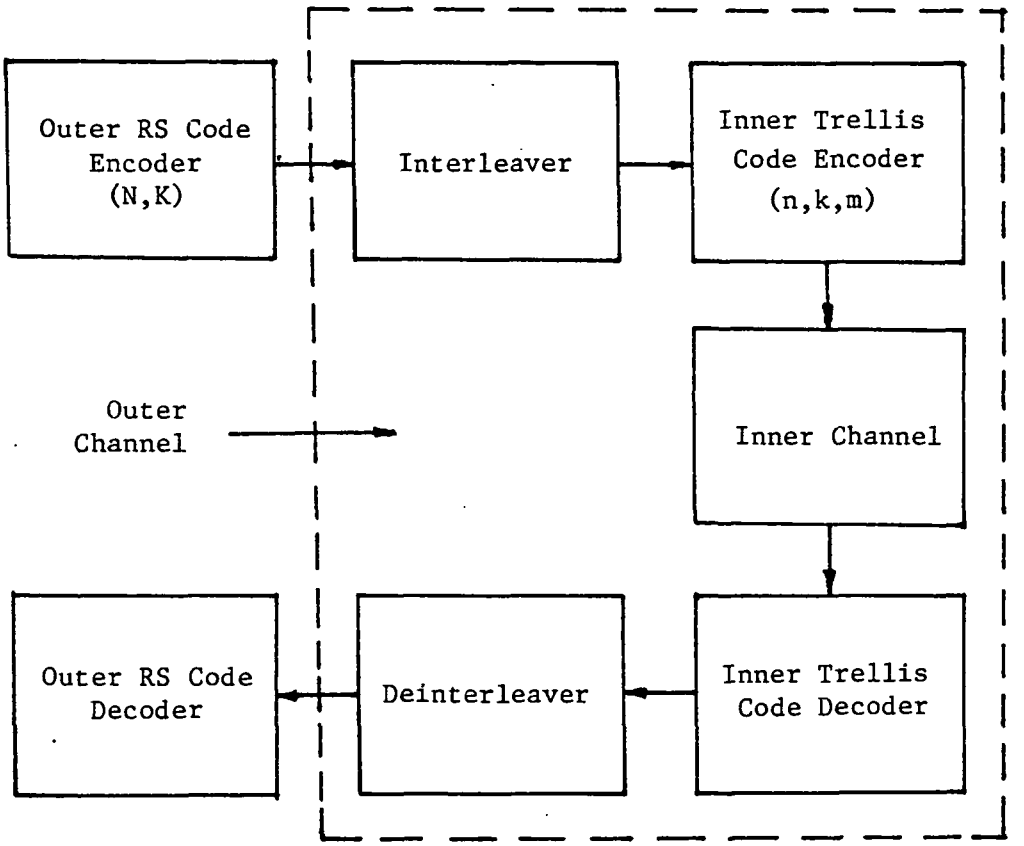

Fig. 1 A concatenated coding system

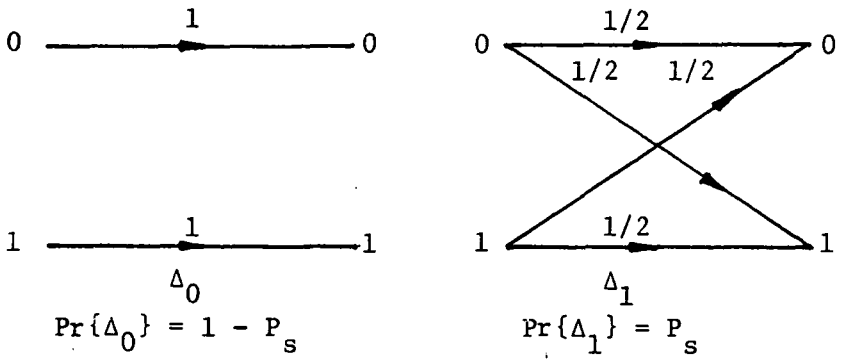

Fig. 2 Outer channel model

$C/C_1, \overline{C}/C_1$

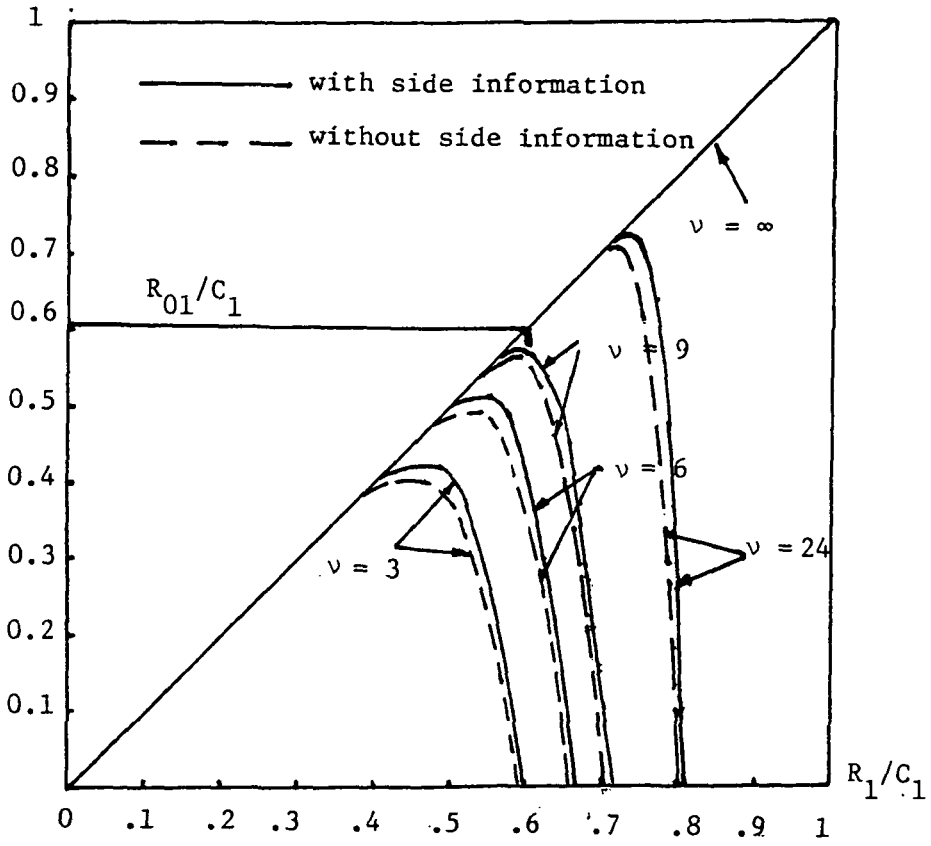

Fig. 3 Capacity for k = 3, b = 8, ε = 0.1.

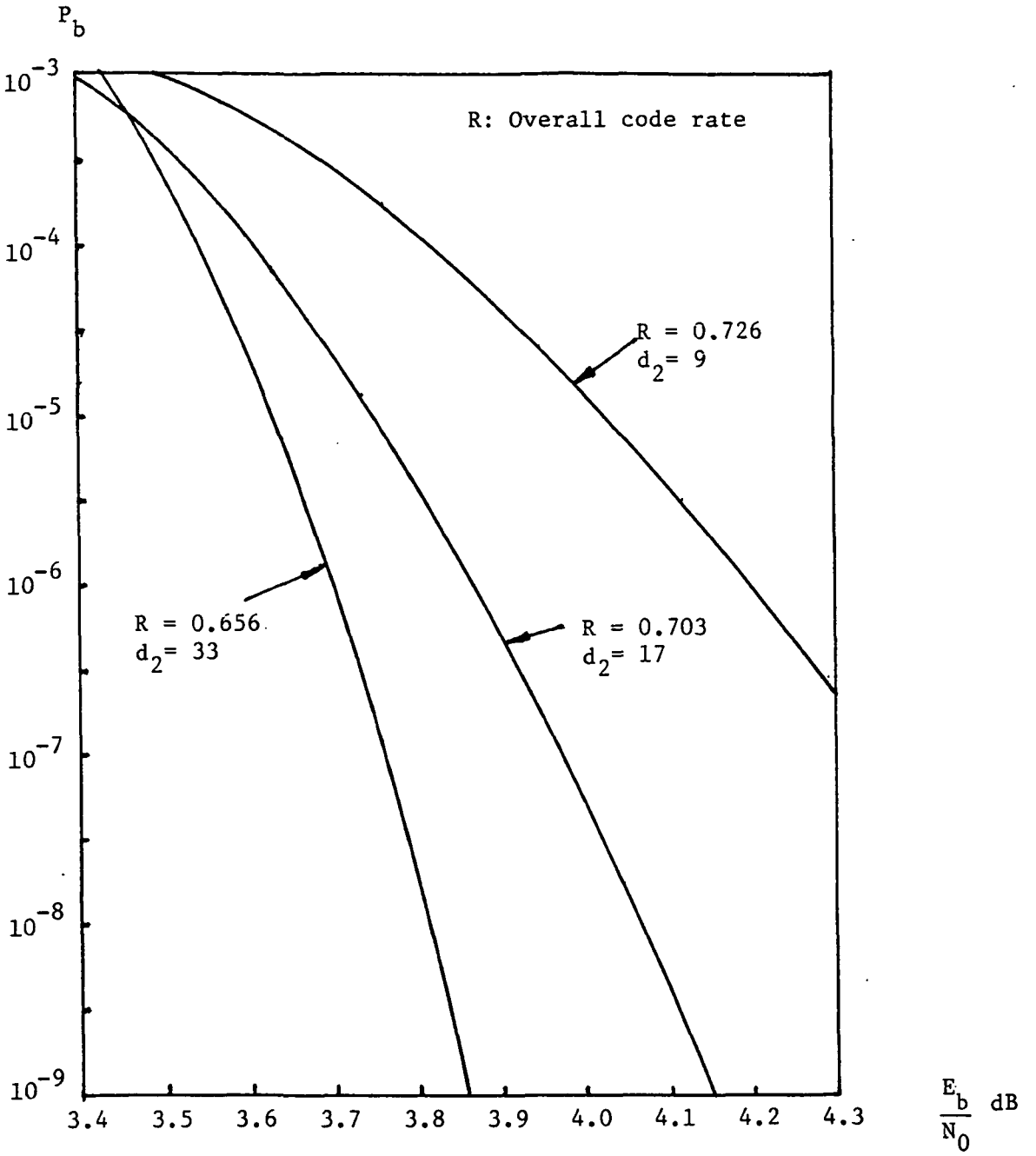

Fig. 4.1 Performance of Example 3.1 (simulation) with R_1 = 3/4 punctured convolutional code as innder code.

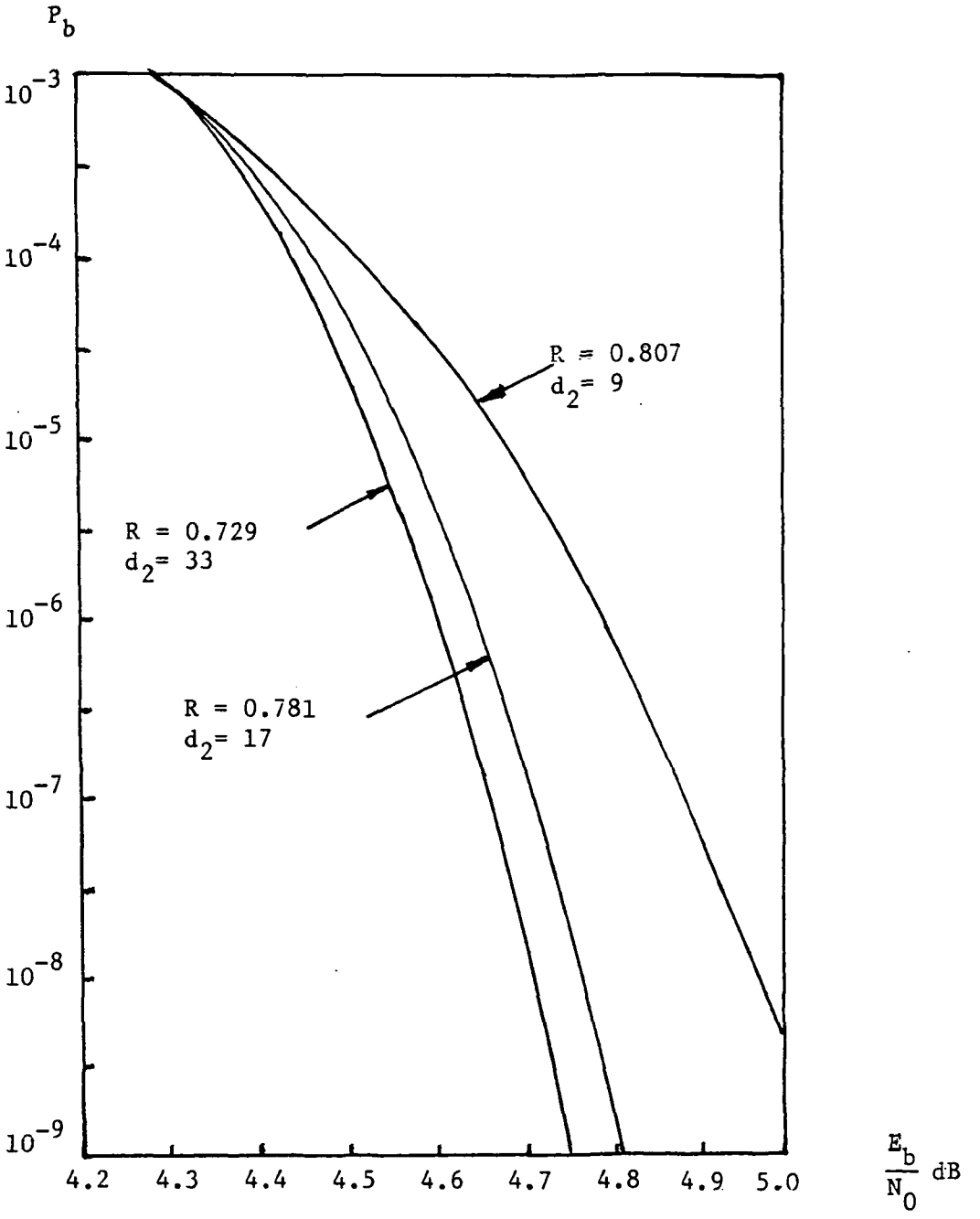

Fig. 4.2 Performance of Example 3.1 (simulation) with R_1 = 5/6 punctured convolutional code as inner code.

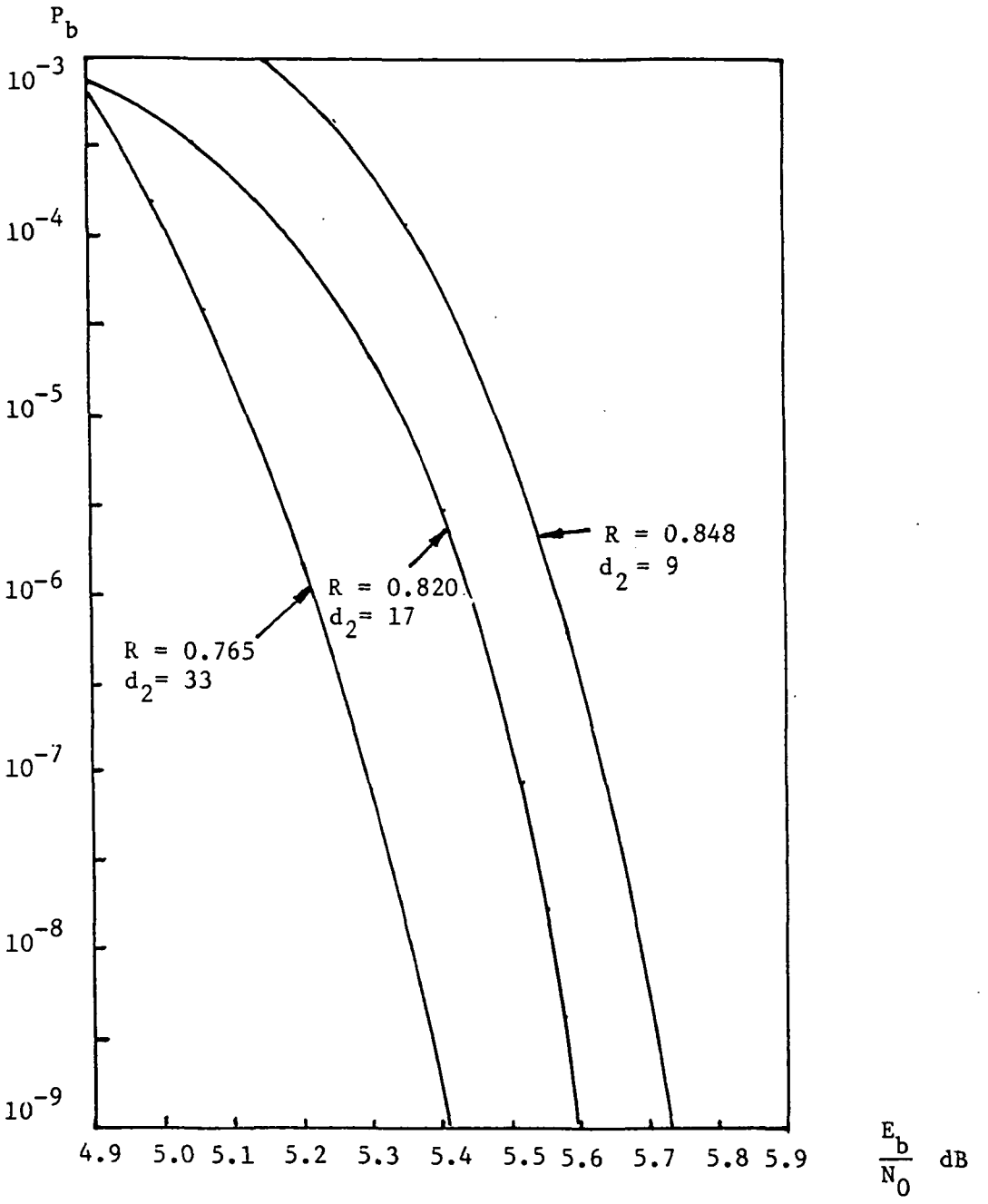

Fig. 4.3 Performance of Example 3.1 (simulation) with R_1 = 7/8 punctured convolutional code as inner code.

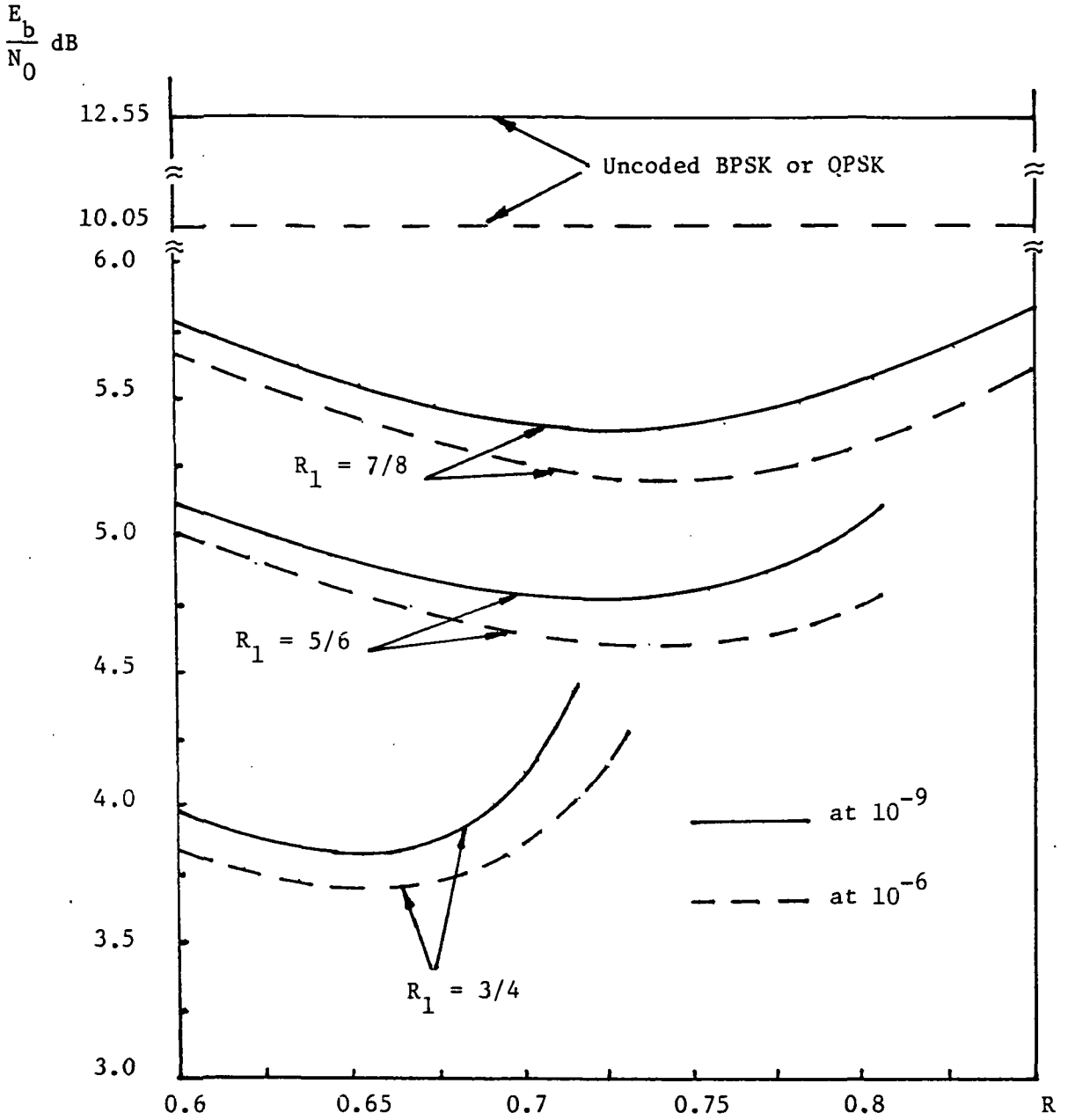

Fig. 4.4 E_b/N_0 required to achieve $P_b = 10^{-6}$ and 10^{-9} vs. overall code rate for Example 3.1 with 64-state punctured convolutional code as inner code (simulation).

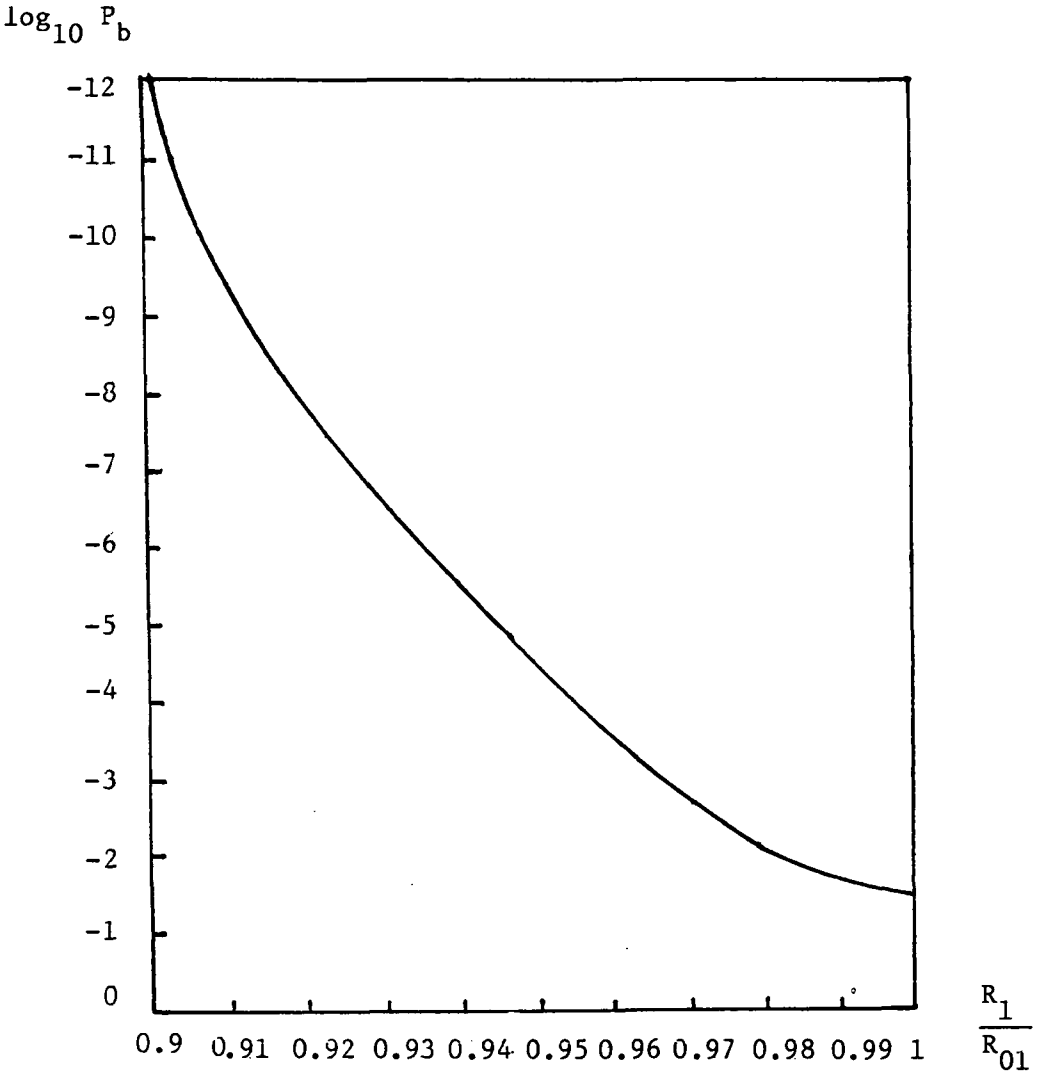

Fig. 4.5 Code performance of Example 3.1 with $R_1 = 3/4$, $d_2 = 33$, $R = 0.656$ vs. R_1/R_{01}.

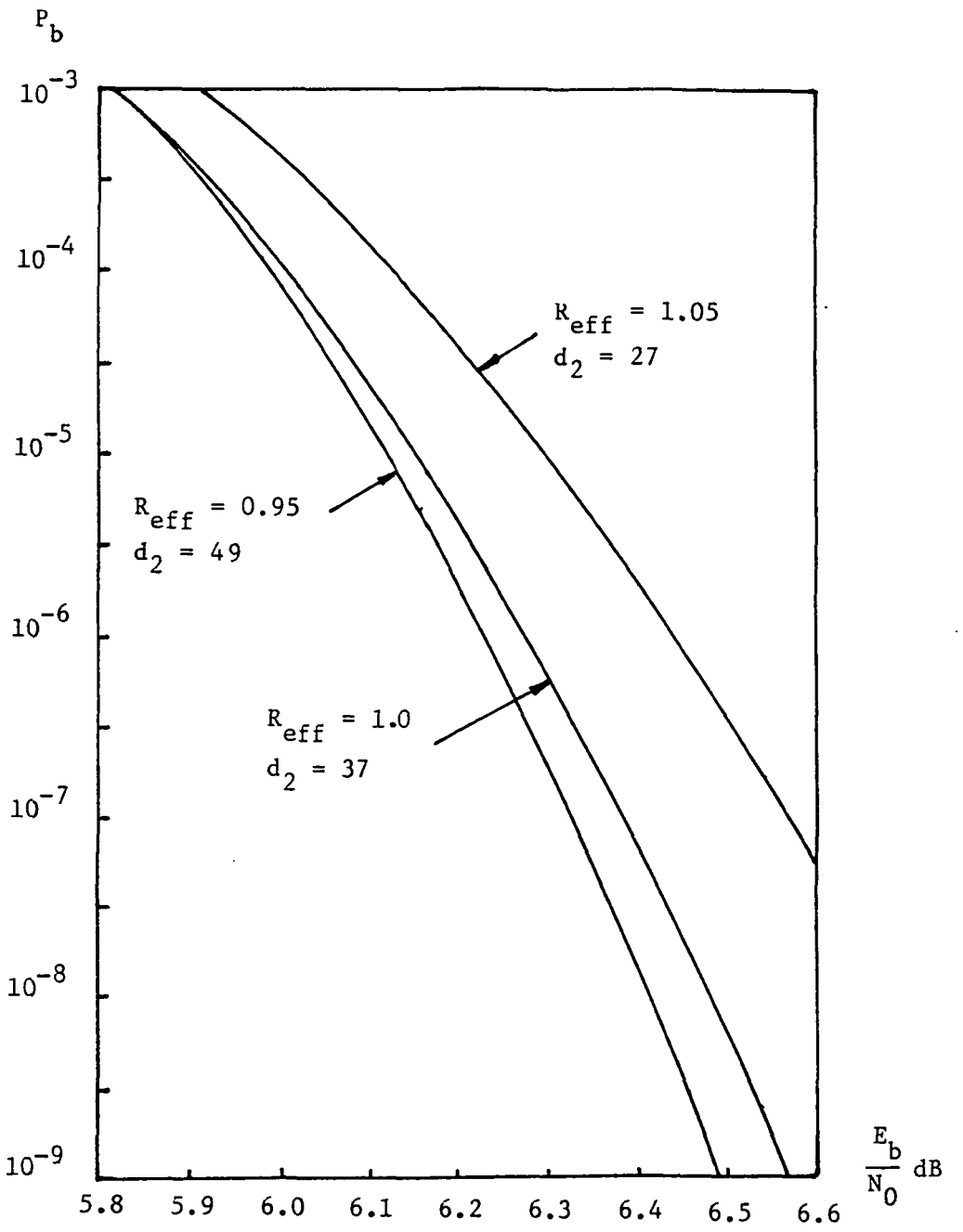

Fig. 5.1 Performance of Example 3.2 (simulation) with 4-state, rate $R_1 = 7/9$ PTVTC as inner code.

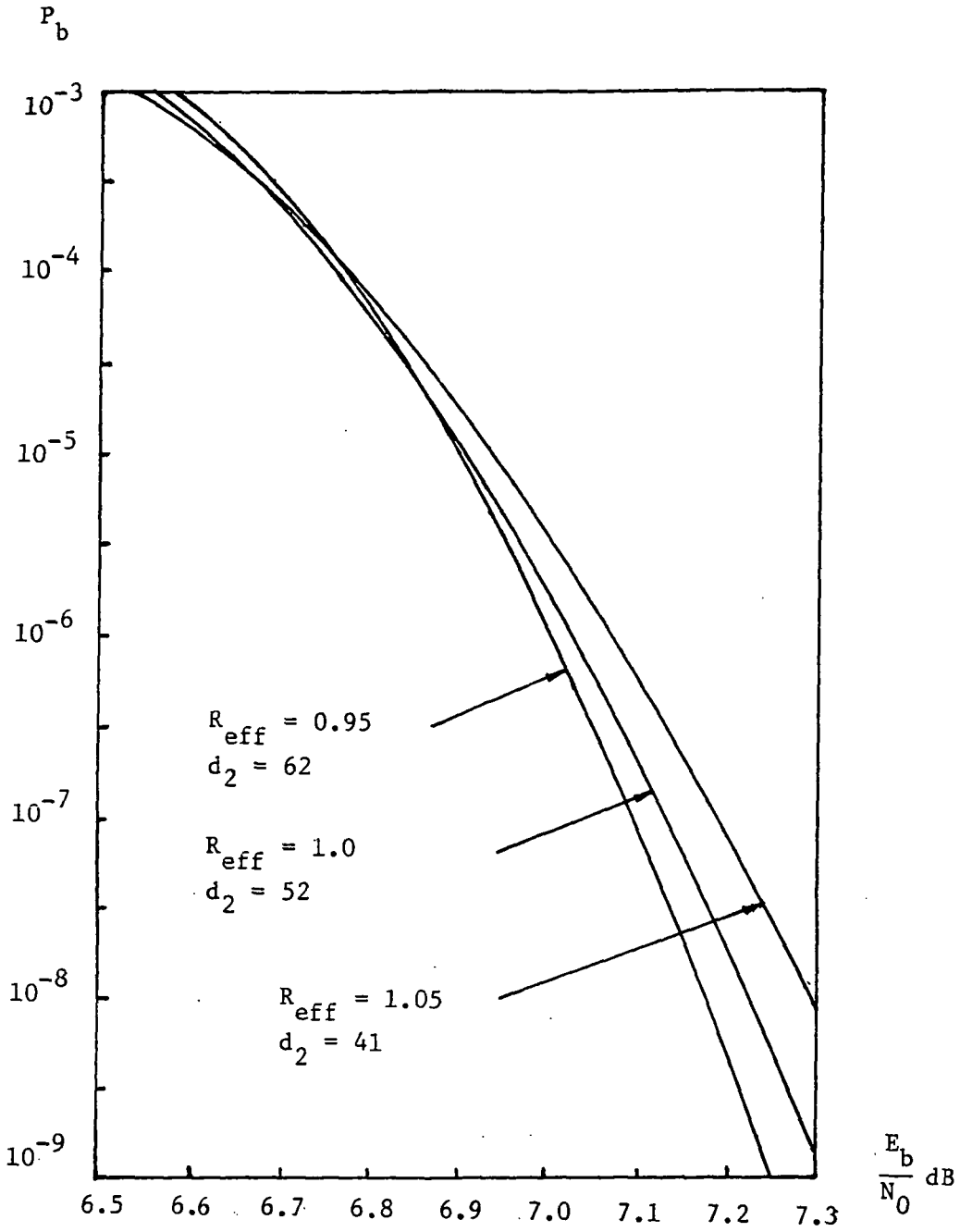

Fig. 5.2 Performance of Example 3.2 (simulation) with 4-state, rate $R_1 = 5/6$ PTVTC as inner code.

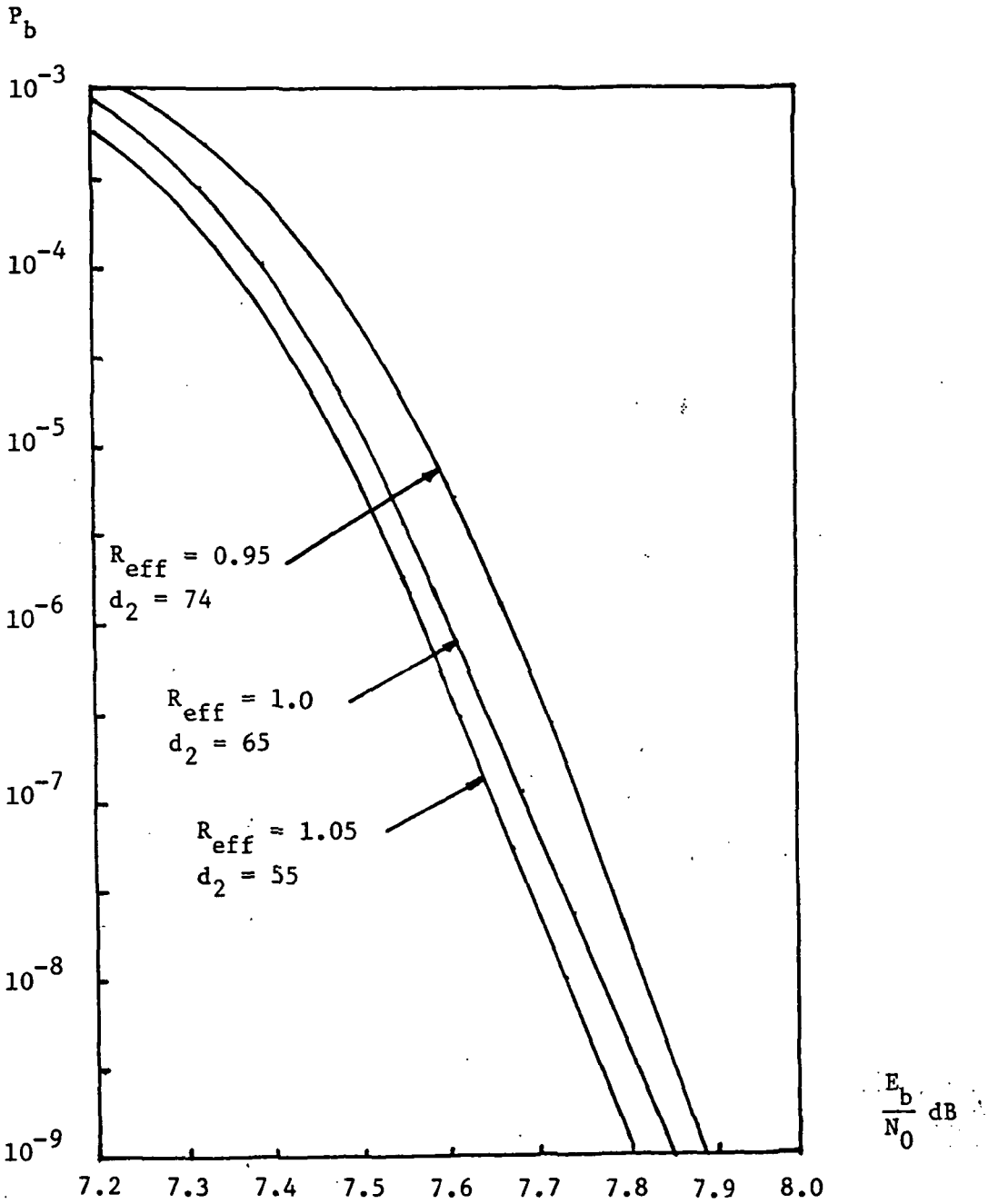

Fig. 5.3 Performance of Example 3.2 (simulation) with 4-state, rate $R_1 = 8/9$ PTVTC as inner code.

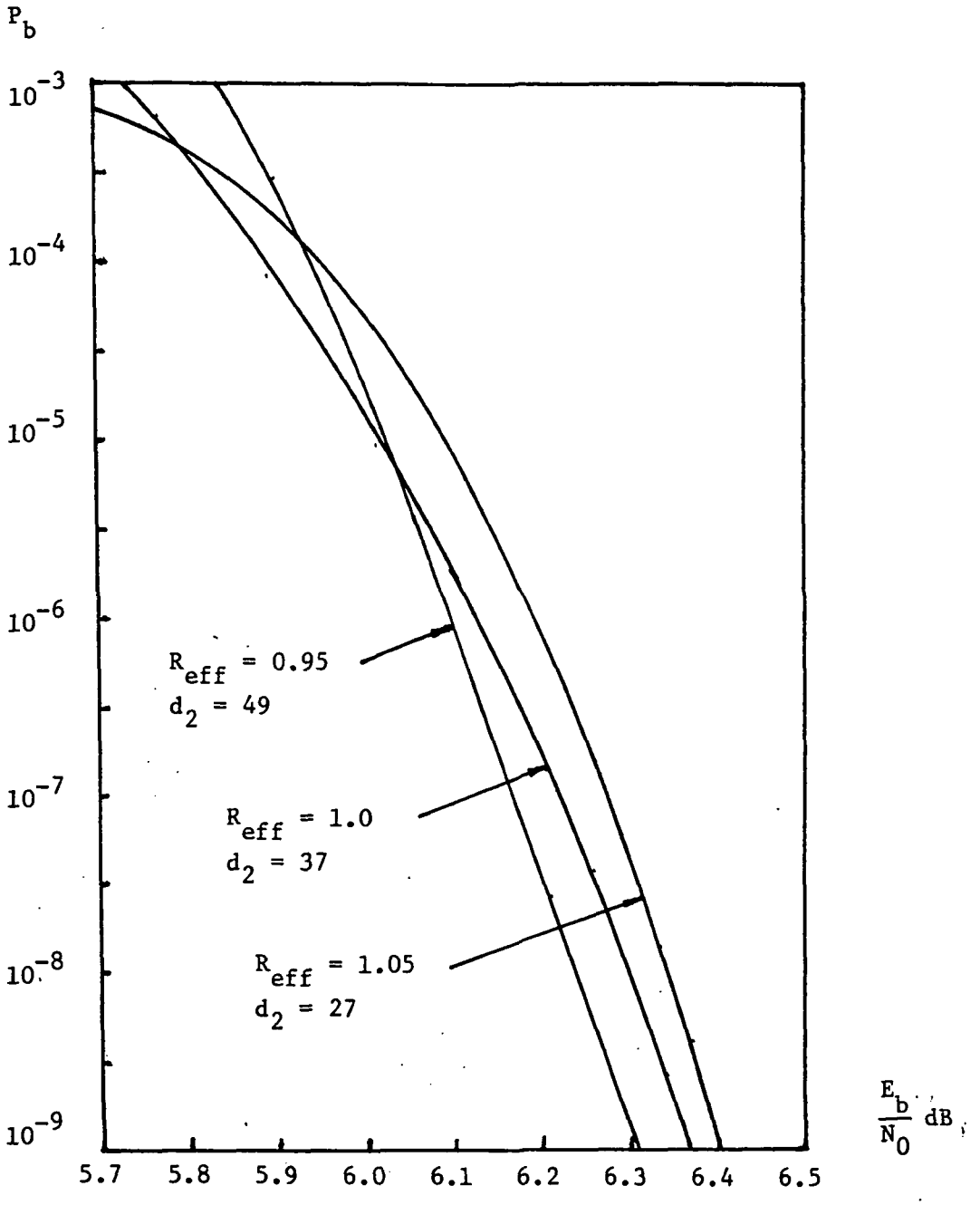

Fig. 5.4 Performance of Example 3.2 (simulation) with 16-state, rate $R_1 = 7/9$ PTVTC as inner code.

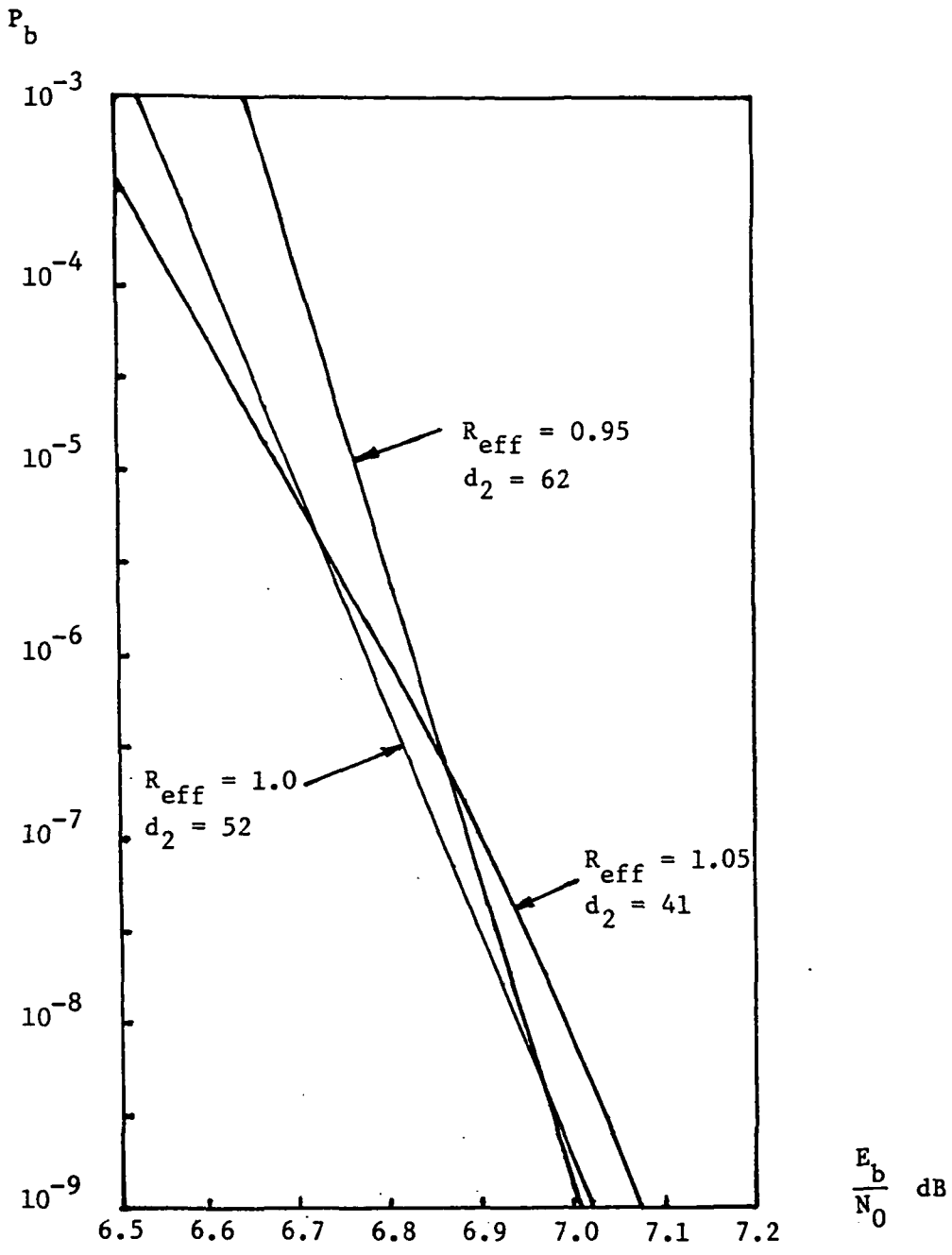

Fig. 5.5 Performance of Example 3.2 (simulation) with 16-state, rate $R_1 = 5/6$ PTVTC as inner code.

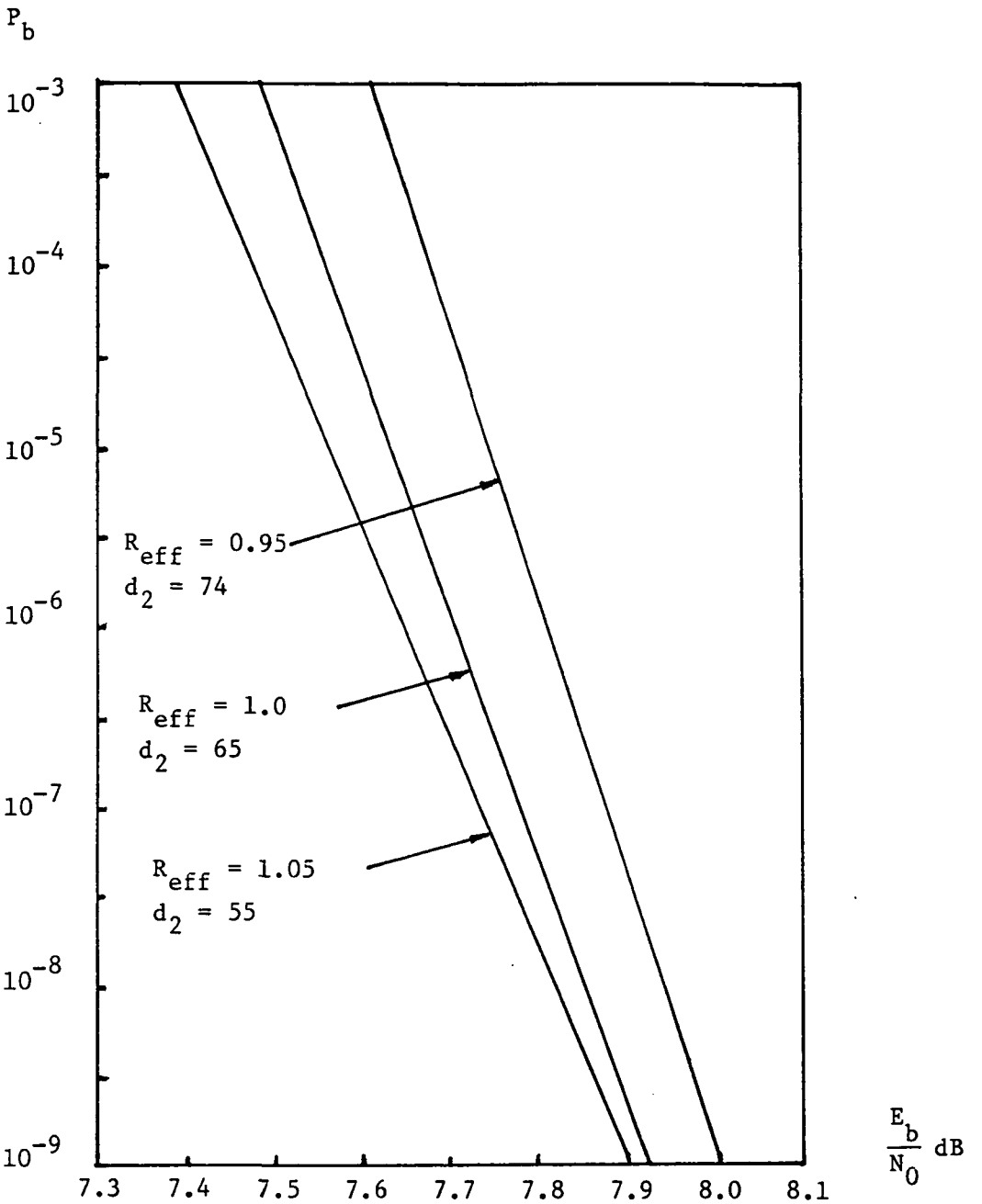

Fig. 5.6 Performance of Example 3.2 (simulation) with 16-state, $R_1 = 8/9$ PTVTC code as inner code.

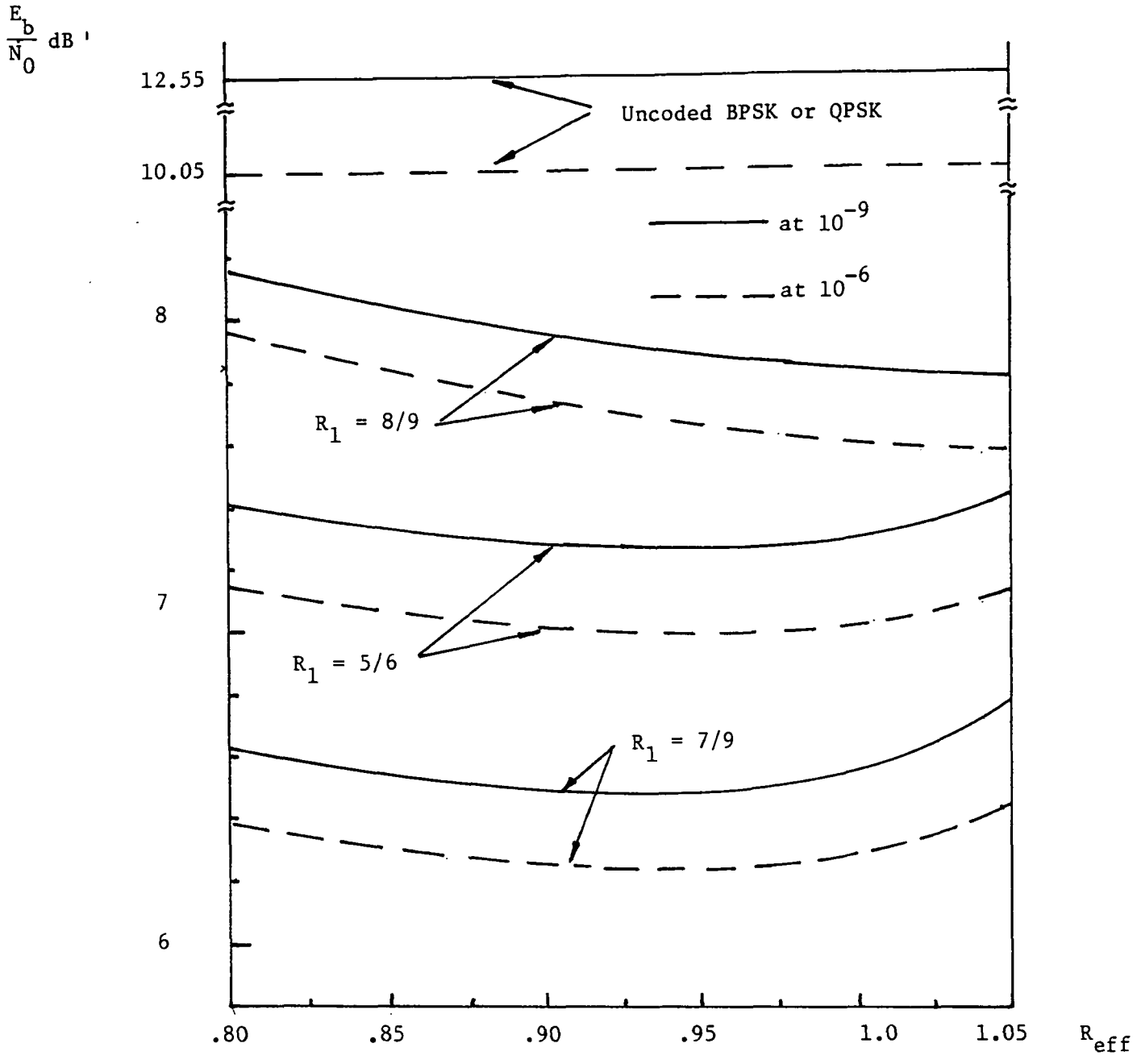

Fig. 5.7 E_b/N_0 required to achieve $P_b = 10^{-6}$ and 10^{-9} vs. R_{eff} for Example 3.2 with 4-state PTVTCs as inner codes (simulation).

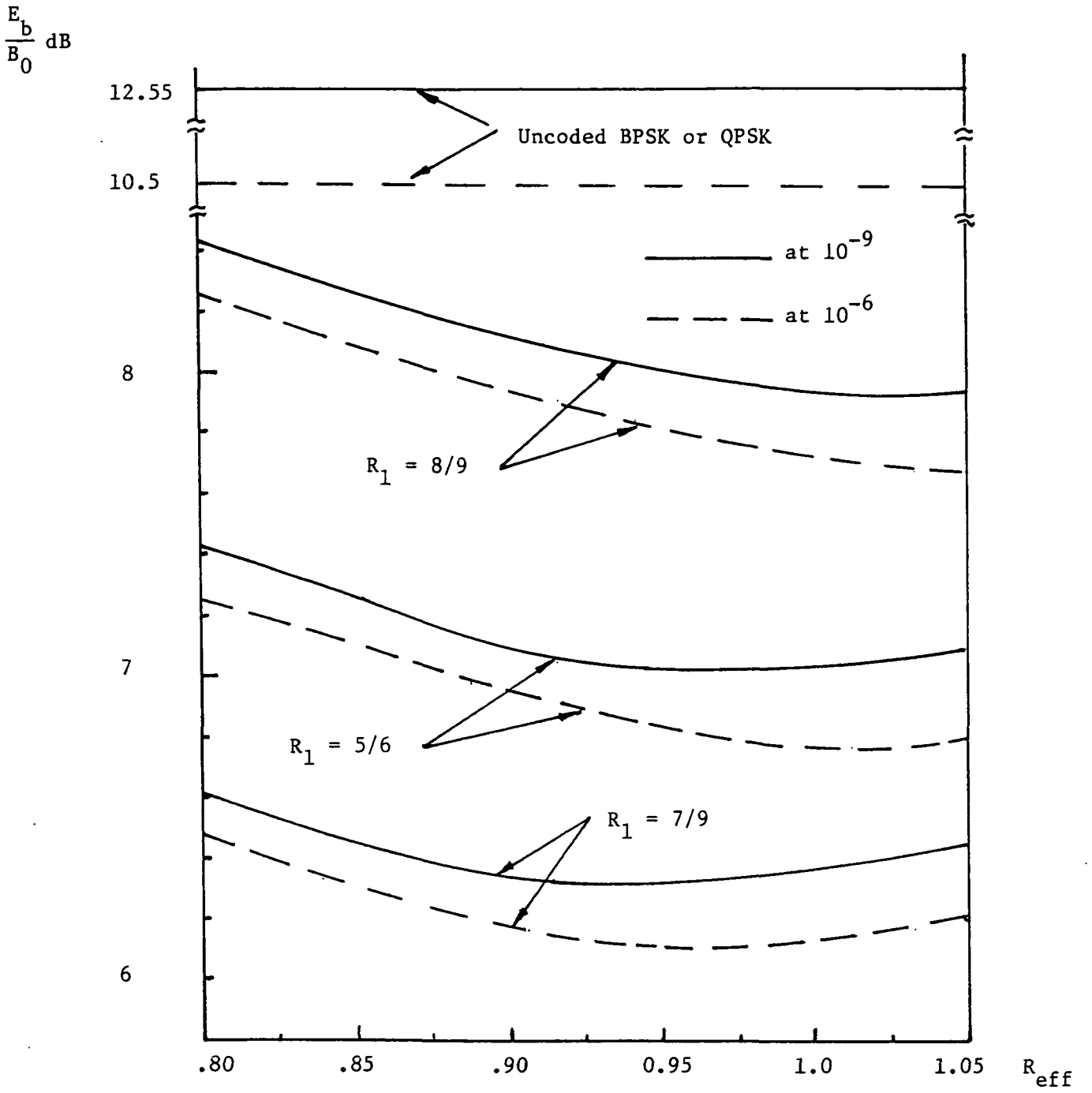

Fig. 5.8 E_b/N_0 required to achieve $P_b = 10^{-6}$ and 10^{-9} vs. R_{eff} for Example 3.2 with 16-state PTVTCs as inner codes (simulation).

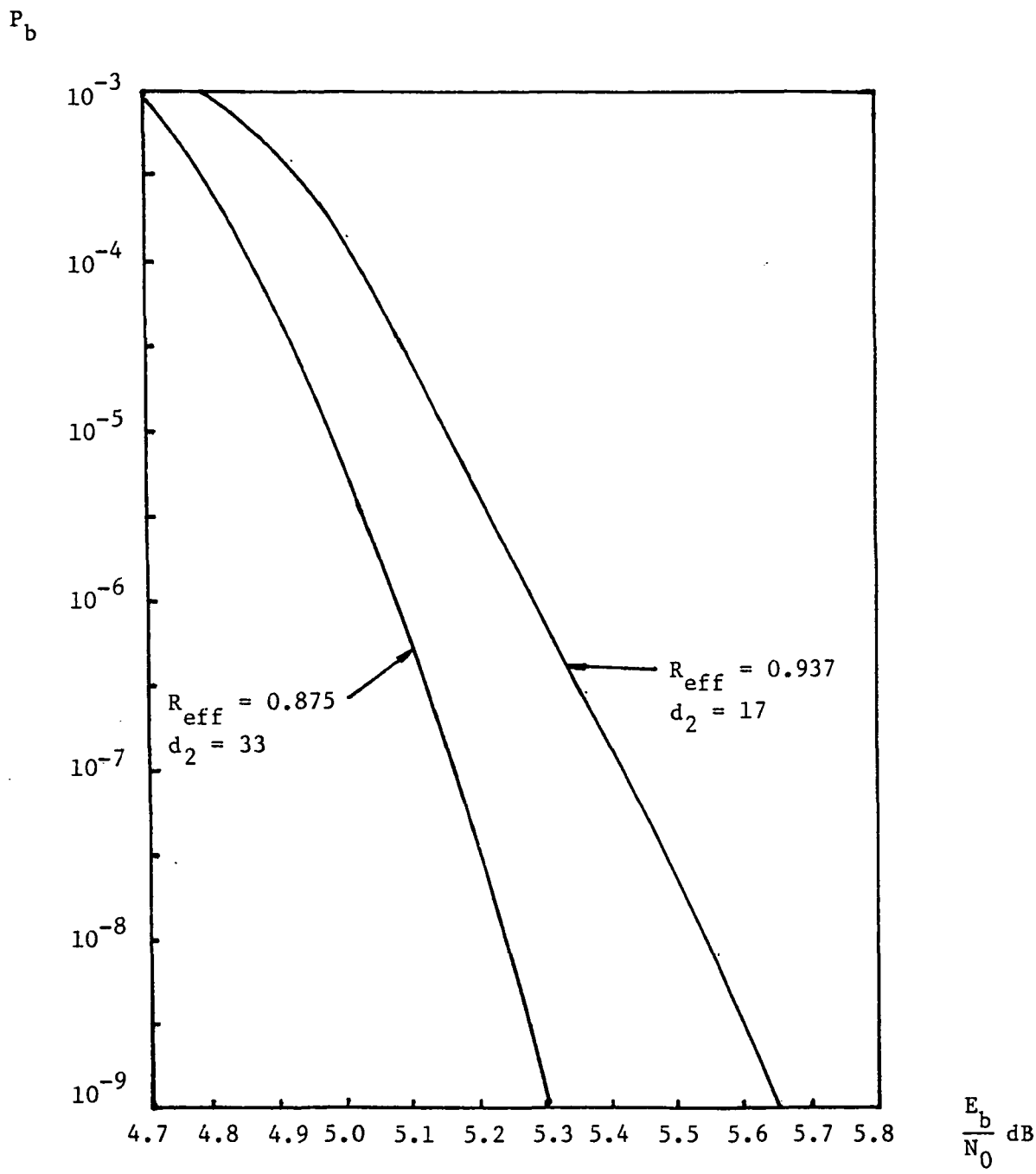

Fig. 6.1 Code performance of Example 3.3 (simulation) with Ungerboeck's 16-state $R_1 = 2/3$ coded 8PSK as inner code.

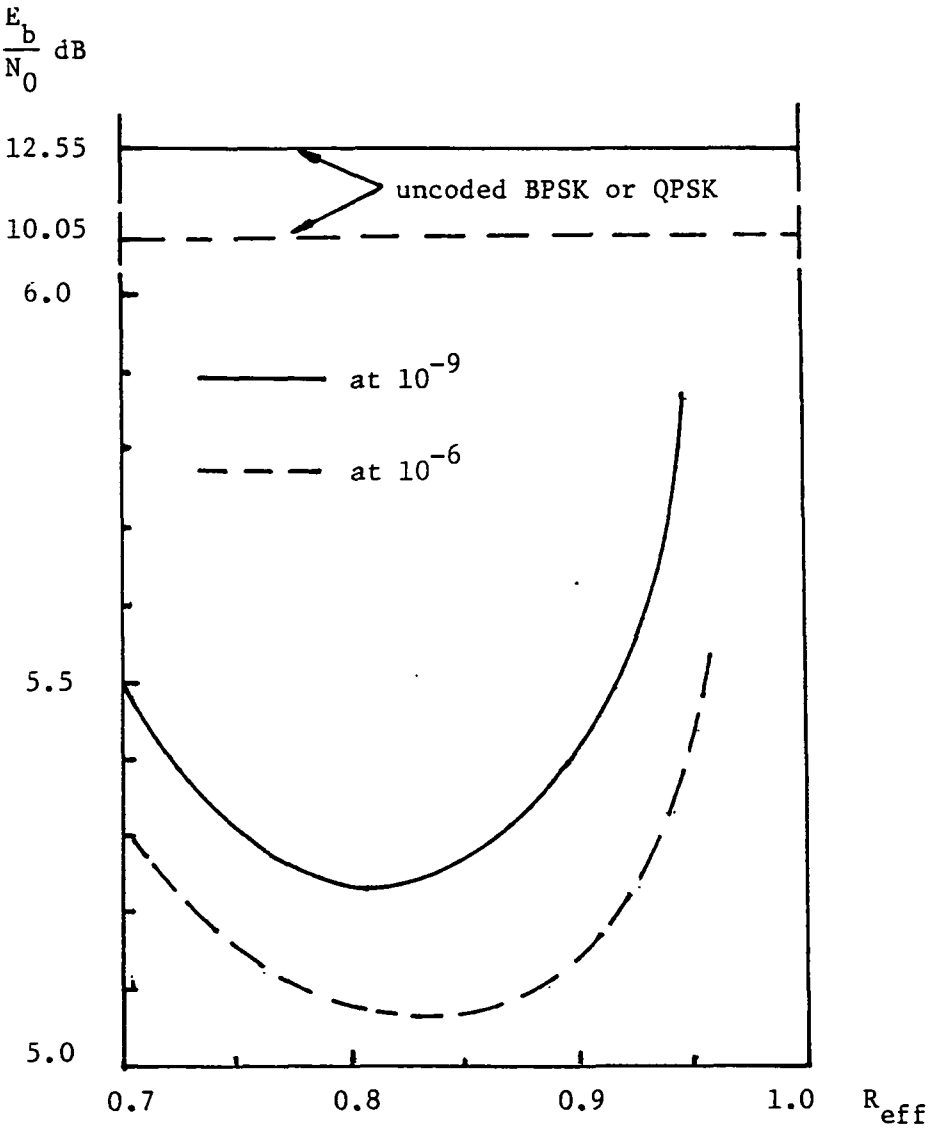

Fig. 6.2 E_b/N_0 required to achieve $P_b = 10^{-6}$ and 10^{-9} vs. the overall effective information rate R_{eff} for Example 3.3 with Ungerboeck's 16-state $R_1 = 2/3$ coded 8PSK as inner code (simulation).

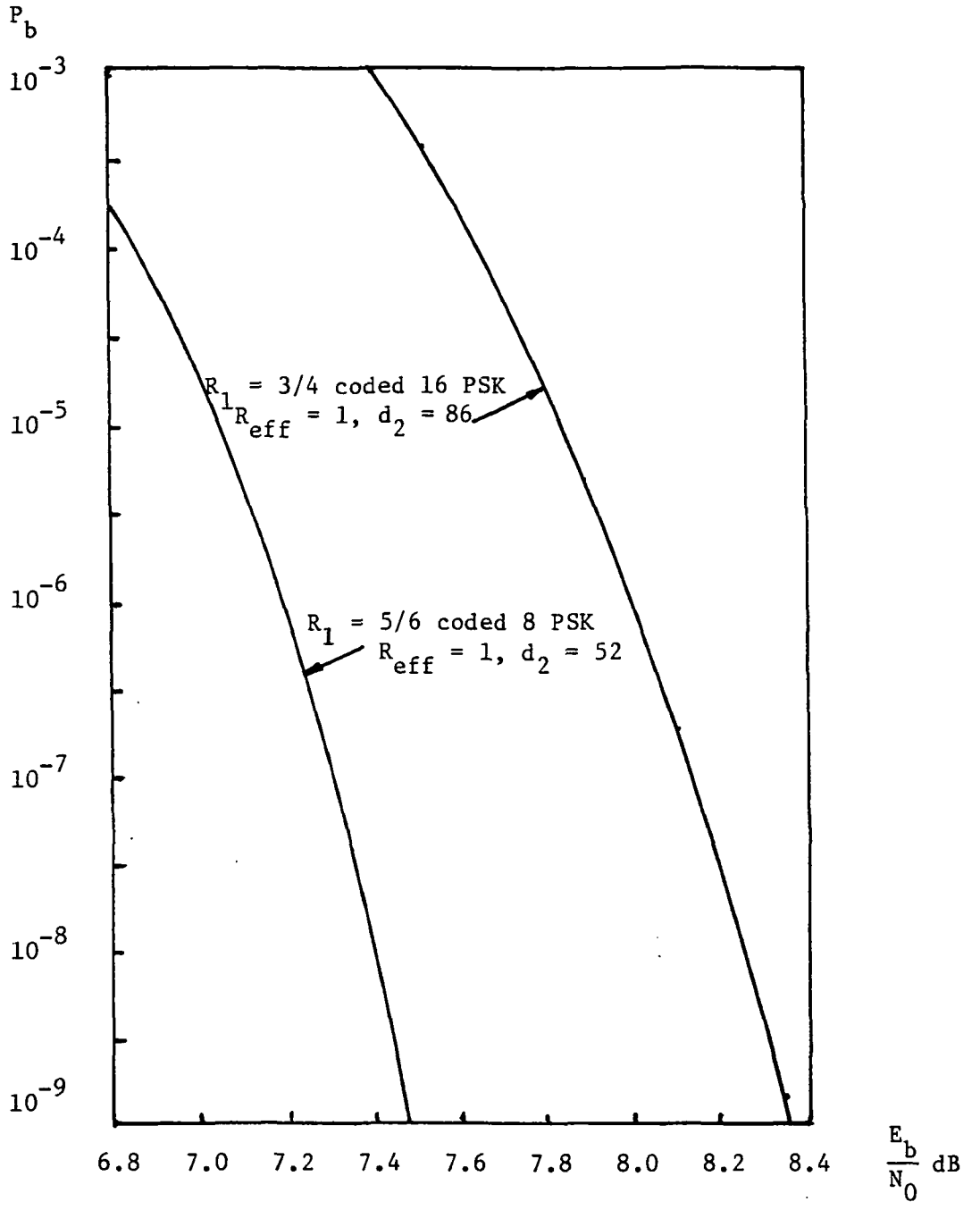

Fig. 7.1 Code performance of Example 3.4 (formula calculation) with 2-state inner codes.

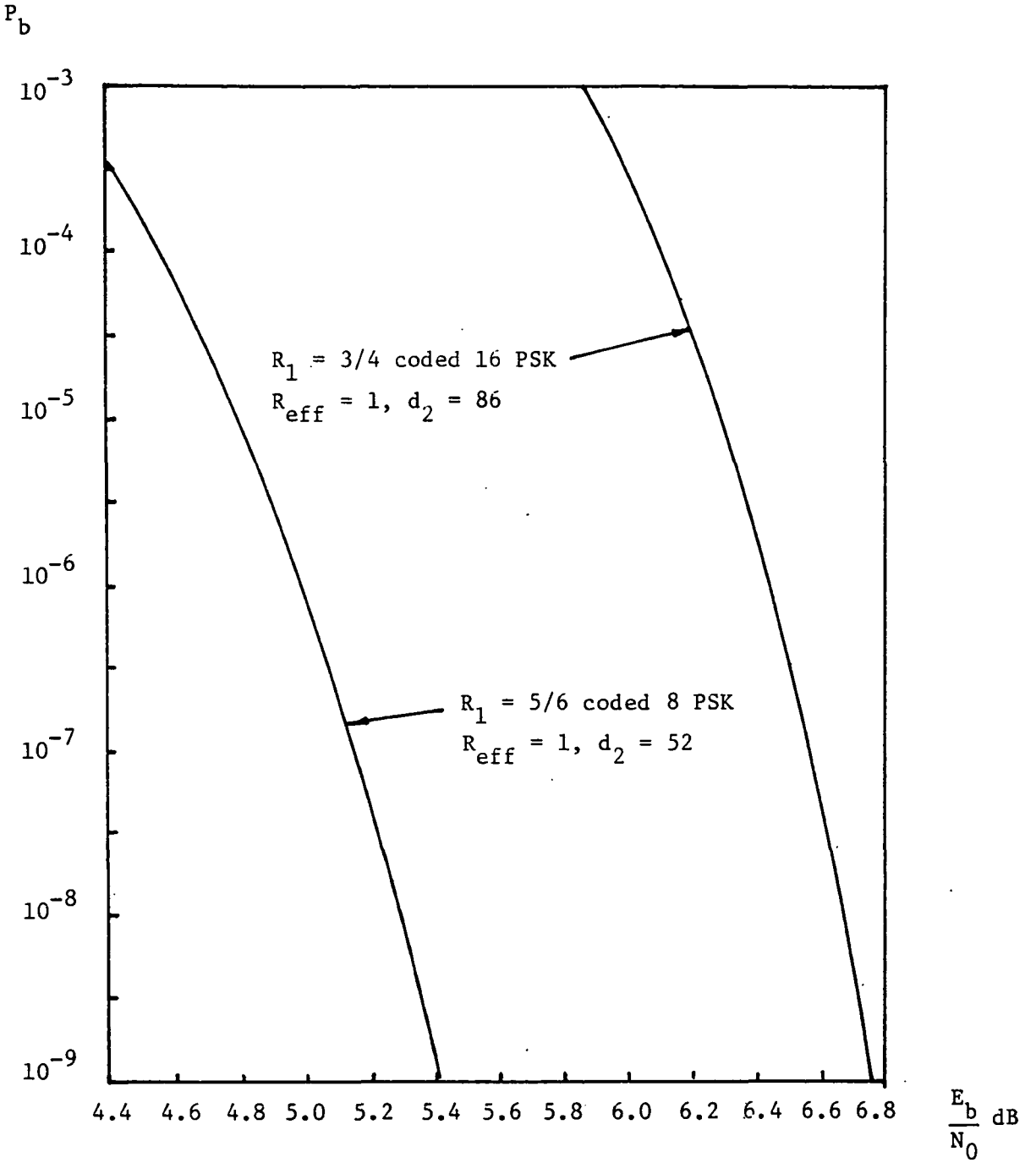

Fig. 7.2 Code performance of Example 3.4 (formula calculation) with 4-state inner codes.

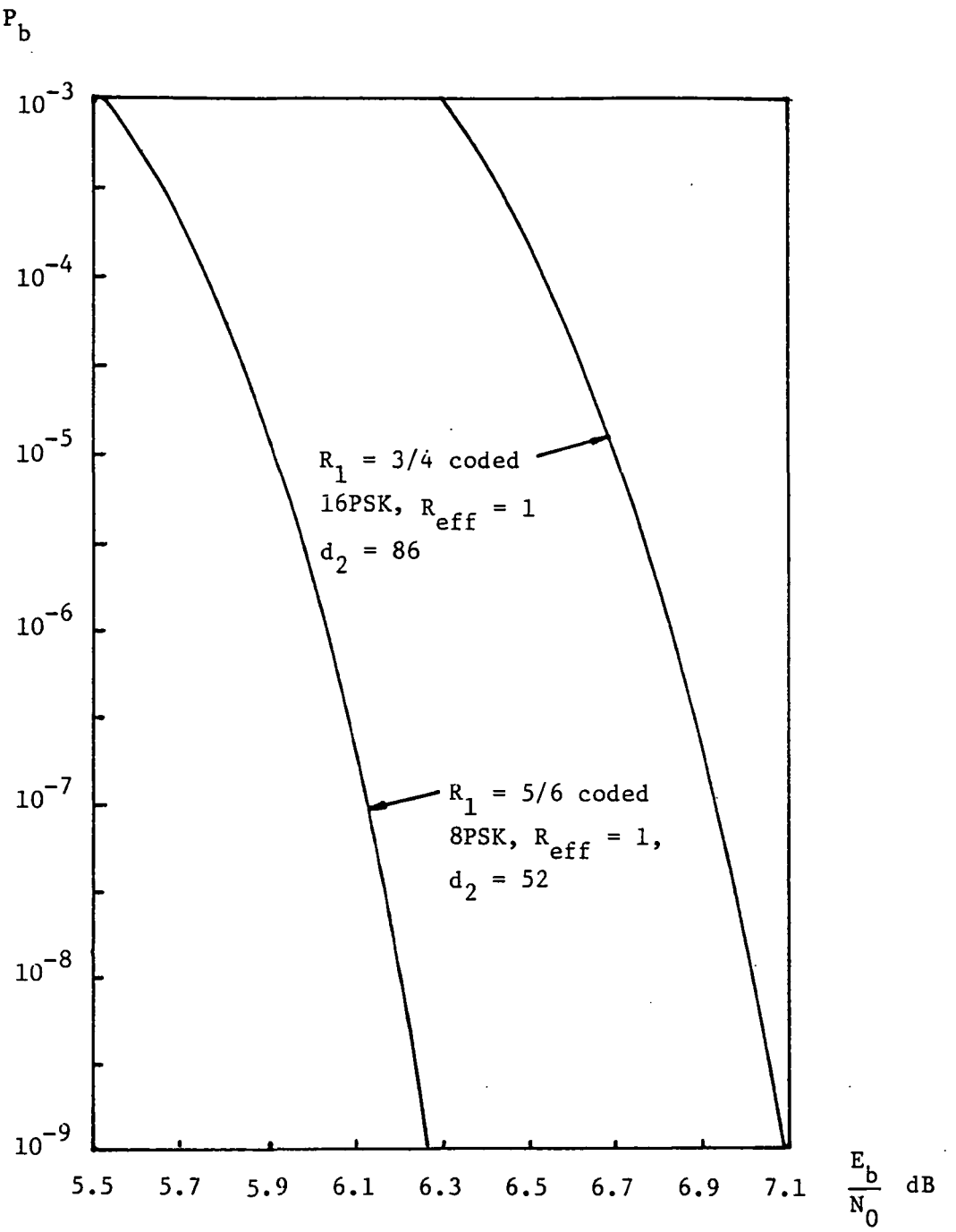

Fig. 7.3 Code performance of Example 3.4 (formula calculation) with 8-state inner codes.

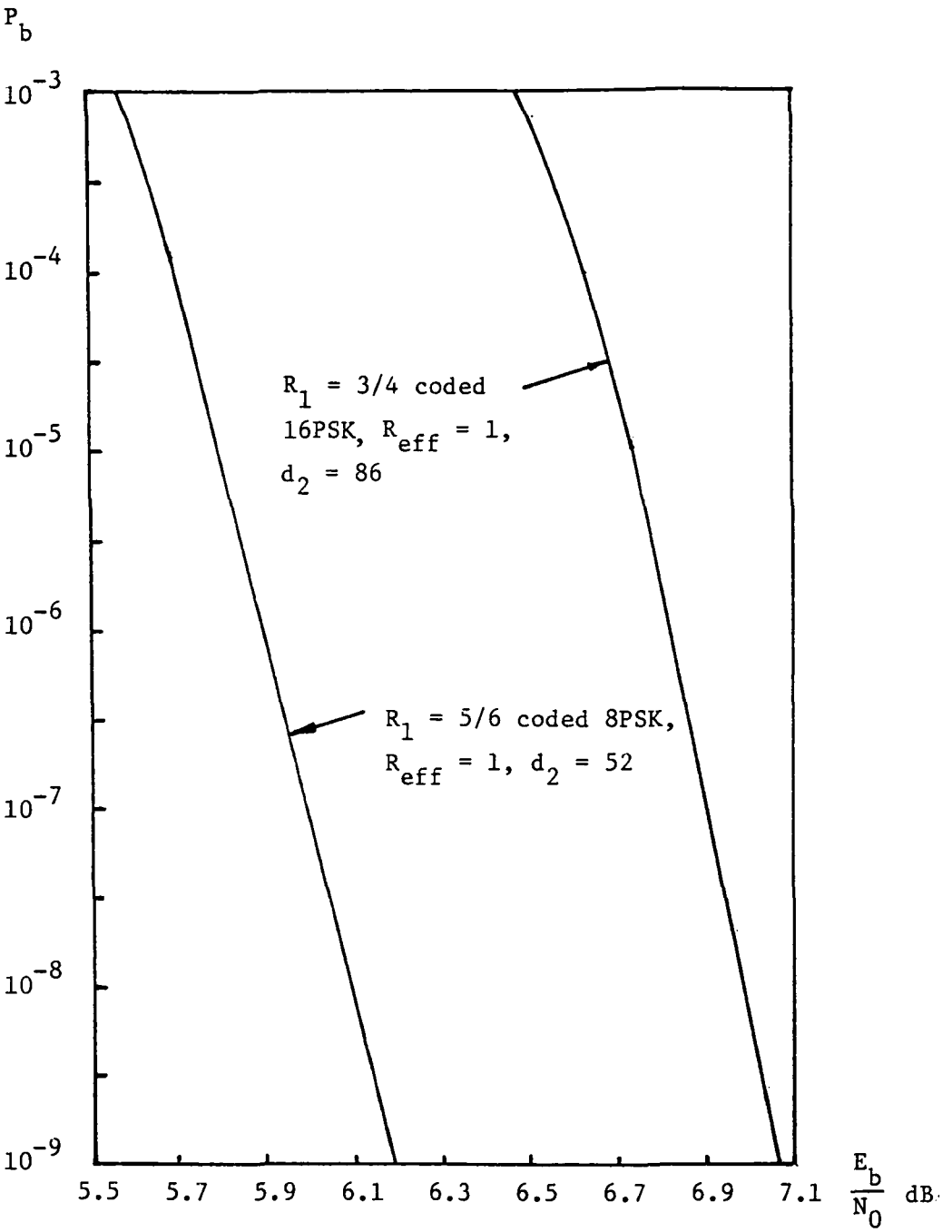

Fig. 7.4 Code performance of Example 3.4 (formula calculation) with 16-state inner codes.

Table 1. Coding gain over QPSK of Example 3.4 with R_1 = 5/6 coded 8PSK as inner code.

# of inner code states	R_{eff}, bits per channel use	Coding gain at 10^{-6} (E_b/N_0)dB, 10^{-6}	Coding gain at 10^{-9} (E_b/N_0)dB, 10^{-9}
2	1	2.85	5.08
4	1	5.07	7.13
8	1	3.65	6.28
16	1	4.15	6.35

Table 2. Coding gain over QPSK of Example 3.4 with R_1 = 3/4 coded 16PSK as inner code.

# of inner code states	R_{eff}, bits per channel use	Coding gain at 10^{-6} (E_b/N_0)dB, 10^{-6}	Coding gain at 10^{-9} (E_b/N_0)dB, 10^{-9}
2	1	2.05	4.18
4	1	3.60	5.77
8	1	3.22	5.45
16	1	3.22	5.47

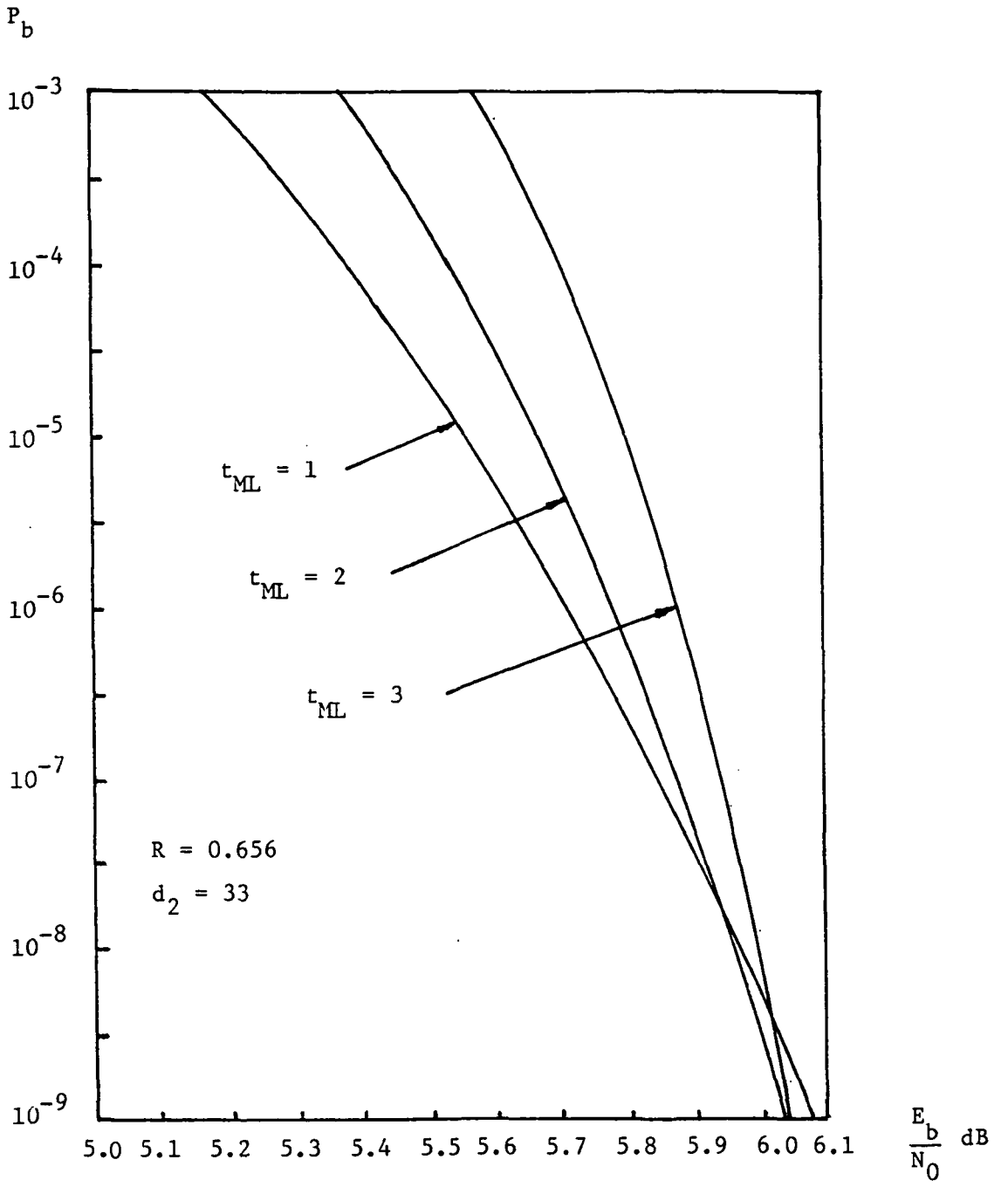

Fig. 8.1 Code performance of Example 3.5 (formula calculation) with $R_1 = 3/4$ self-orthogonal convolutional codes as inner codes.

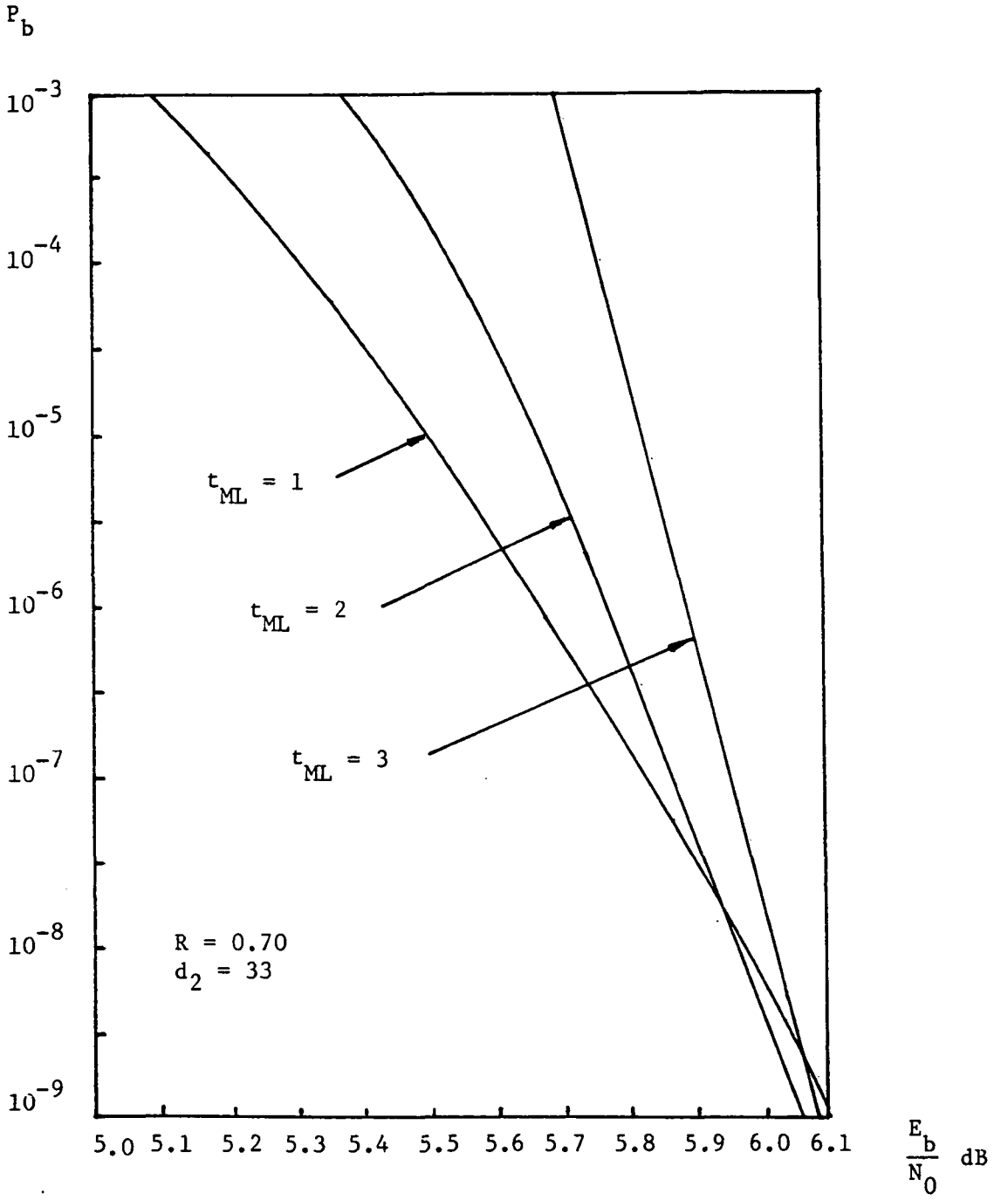

Fig. 8.2 Code performance of Example 3.5 (formula calculation) with $R_1 = 4/5$ self-orthogonal convolutional codes as inner codes.

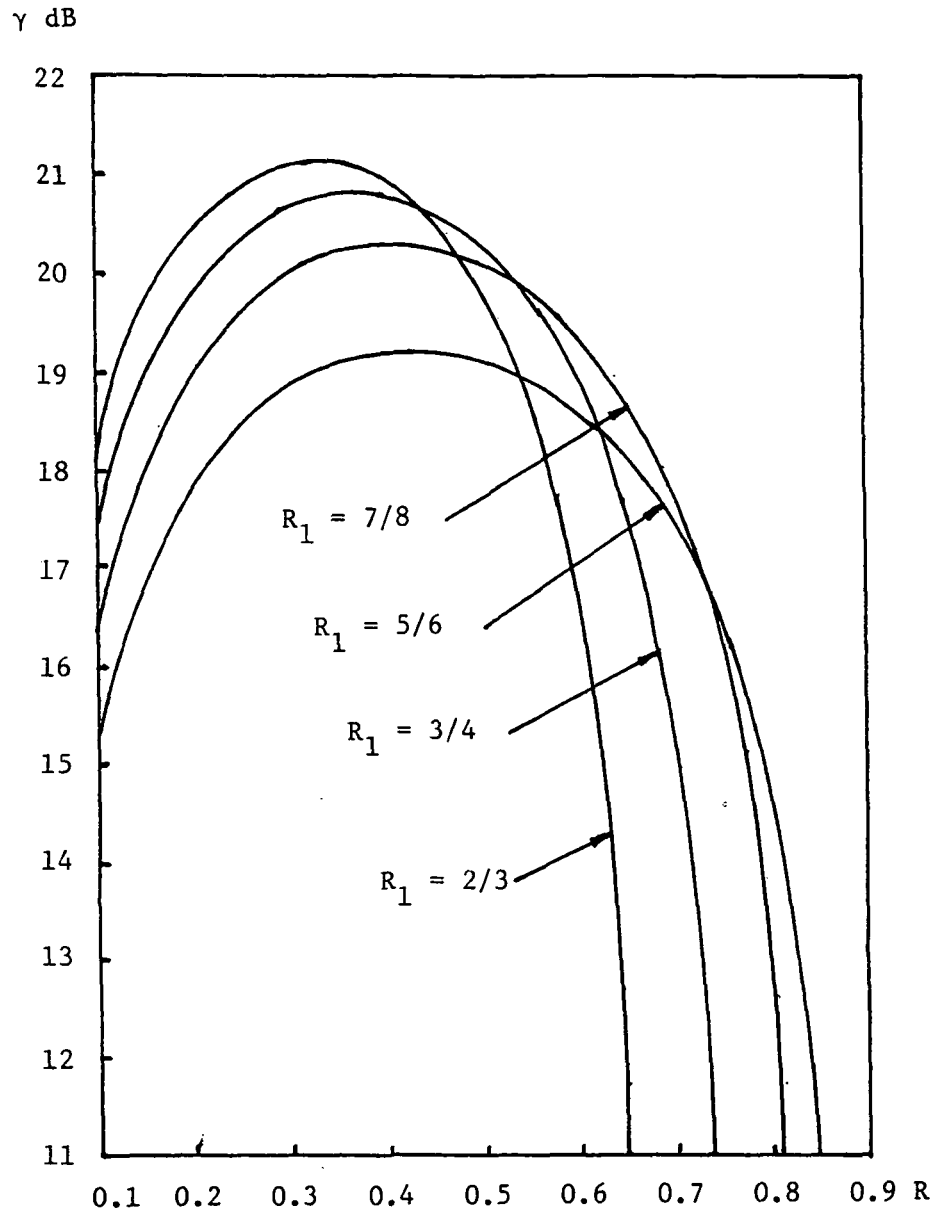

Fig. 9 Asympototic concatenated coding gain without demodulator output quantization and with $R_1 = (n-1)/n$ punctured codes formed from (2,1,6) convolutional code and RS code of length N = 255.

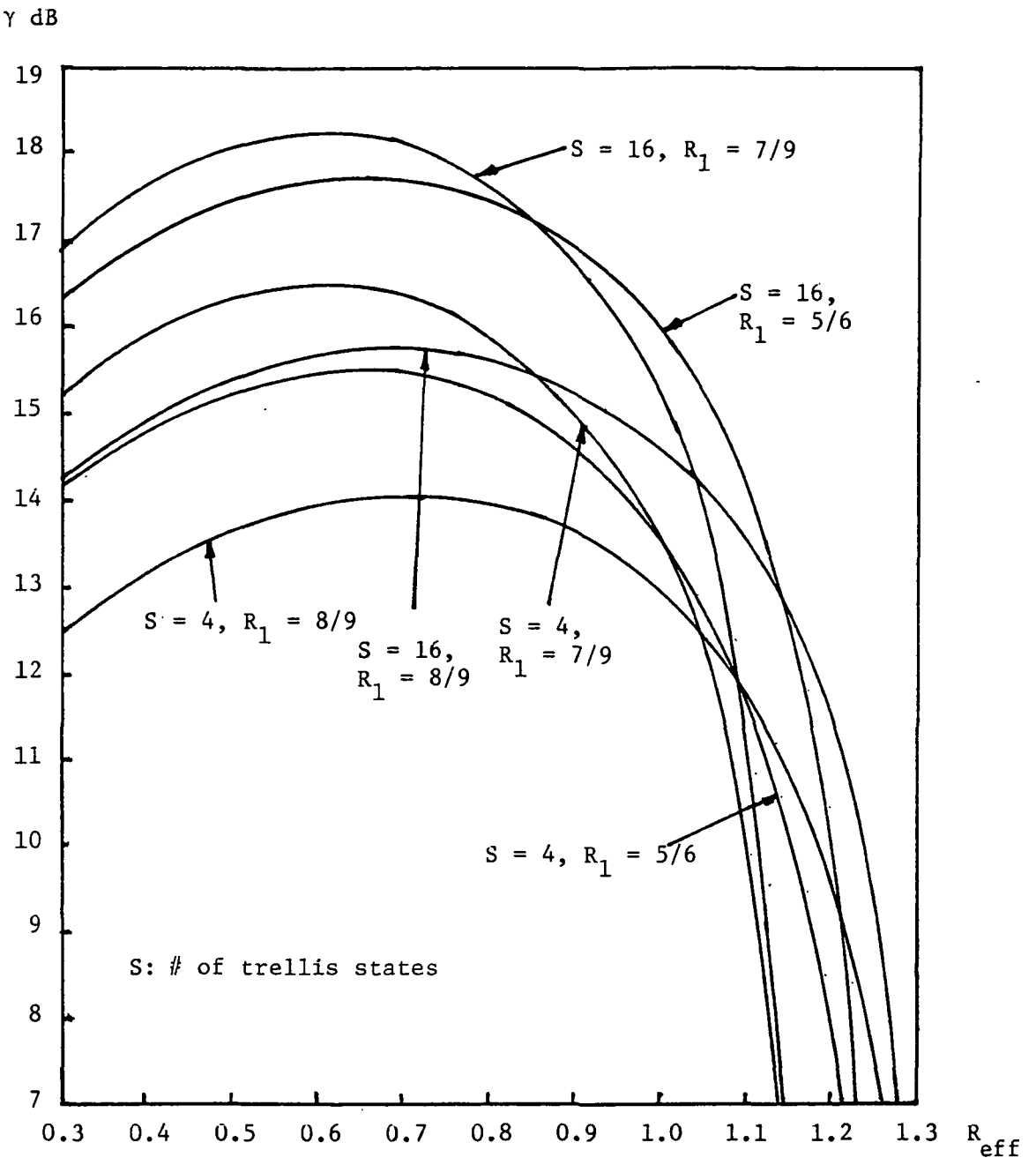

Fig. 10 Asymptotic concatenated coding gain of Example 3.2 with PTVTCs as inner codes and N = 255 RS codes.

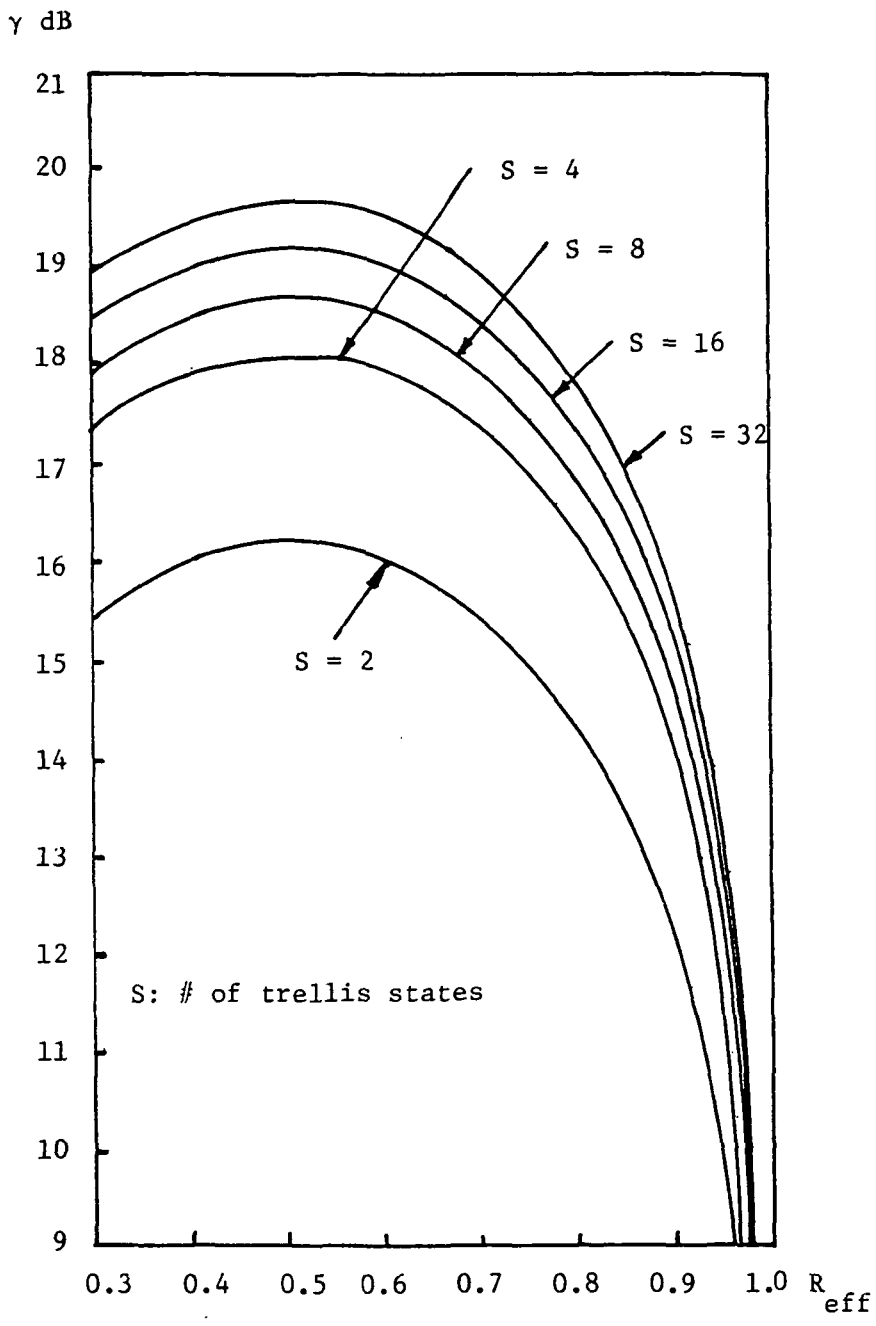

Fig. 11 Asymptotic concatenated coding gain of Example 3.3 with Ungerboeck's $R_1 = 2/3$ coded 8PSK as inner codes and N = 255 RS outer codes.

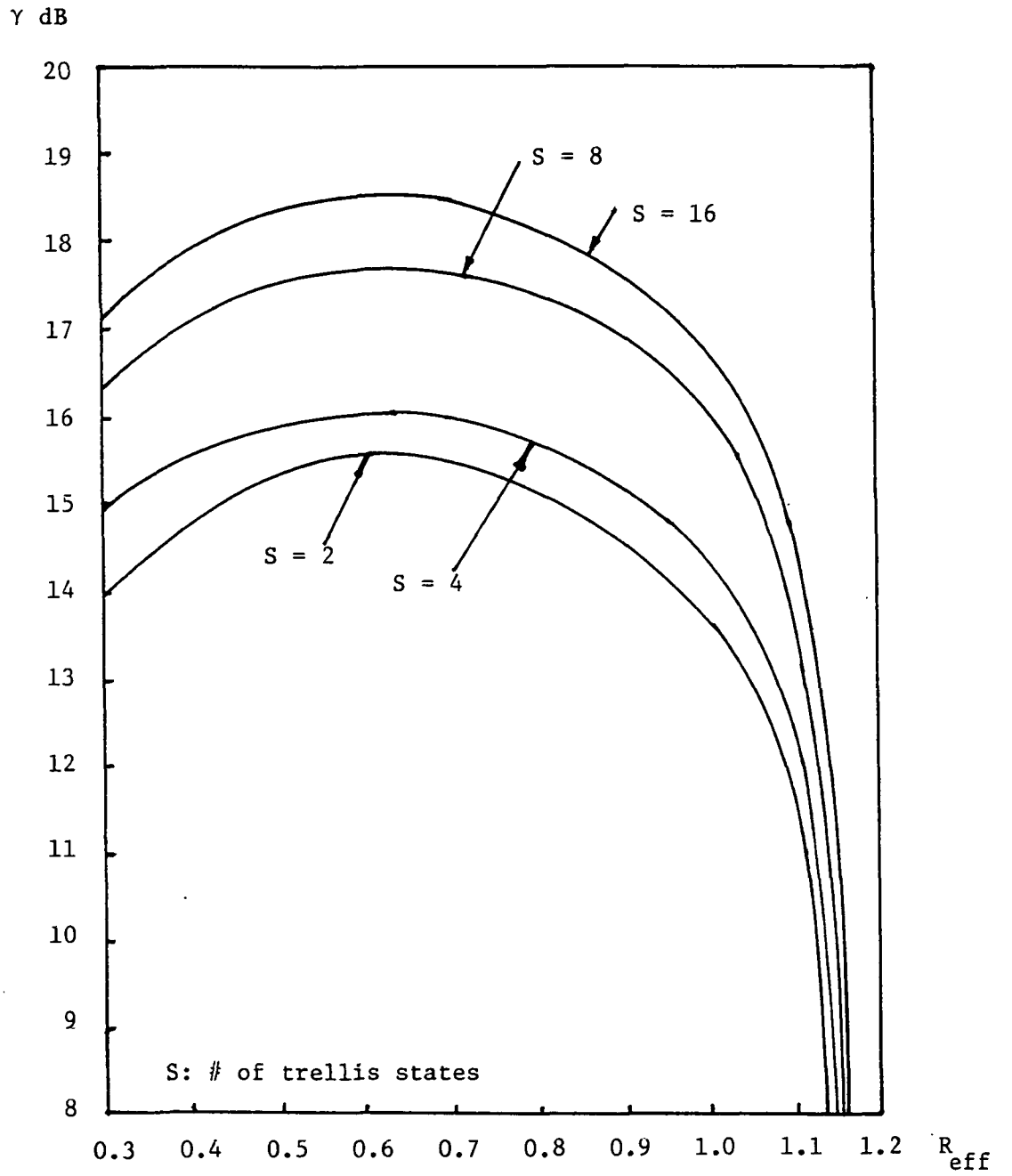

Fig. 12.1 Asymptotic concatenated coding gain of Example 3.4 with $R_1 = 5/6$ coded 8PSK as inner code and N = 255 RS outer codes.

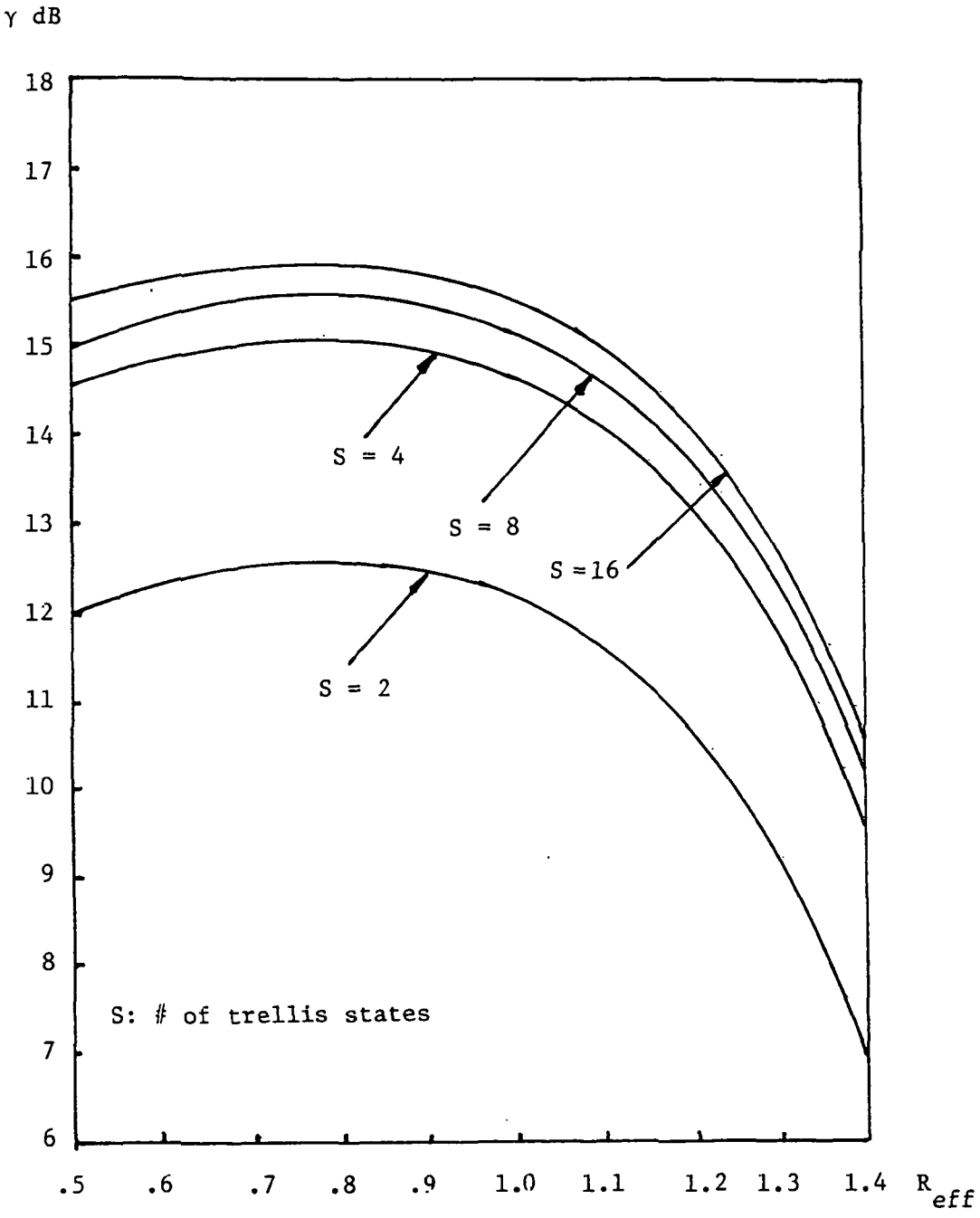

Fig. 12.2 Asymptotic concatenated coding gain of Example 3.4 with R_1 = 3/4 coded 16PSK as inner code and N = 255 RS outer codes.

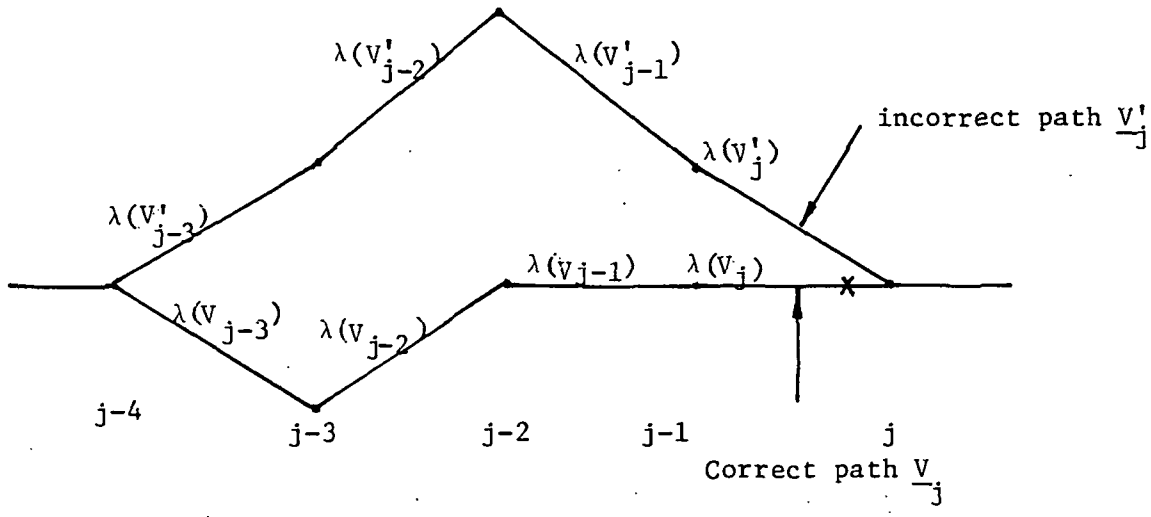

Fig. 13. A 4-branch first event error at level j in decoding a trellis code.

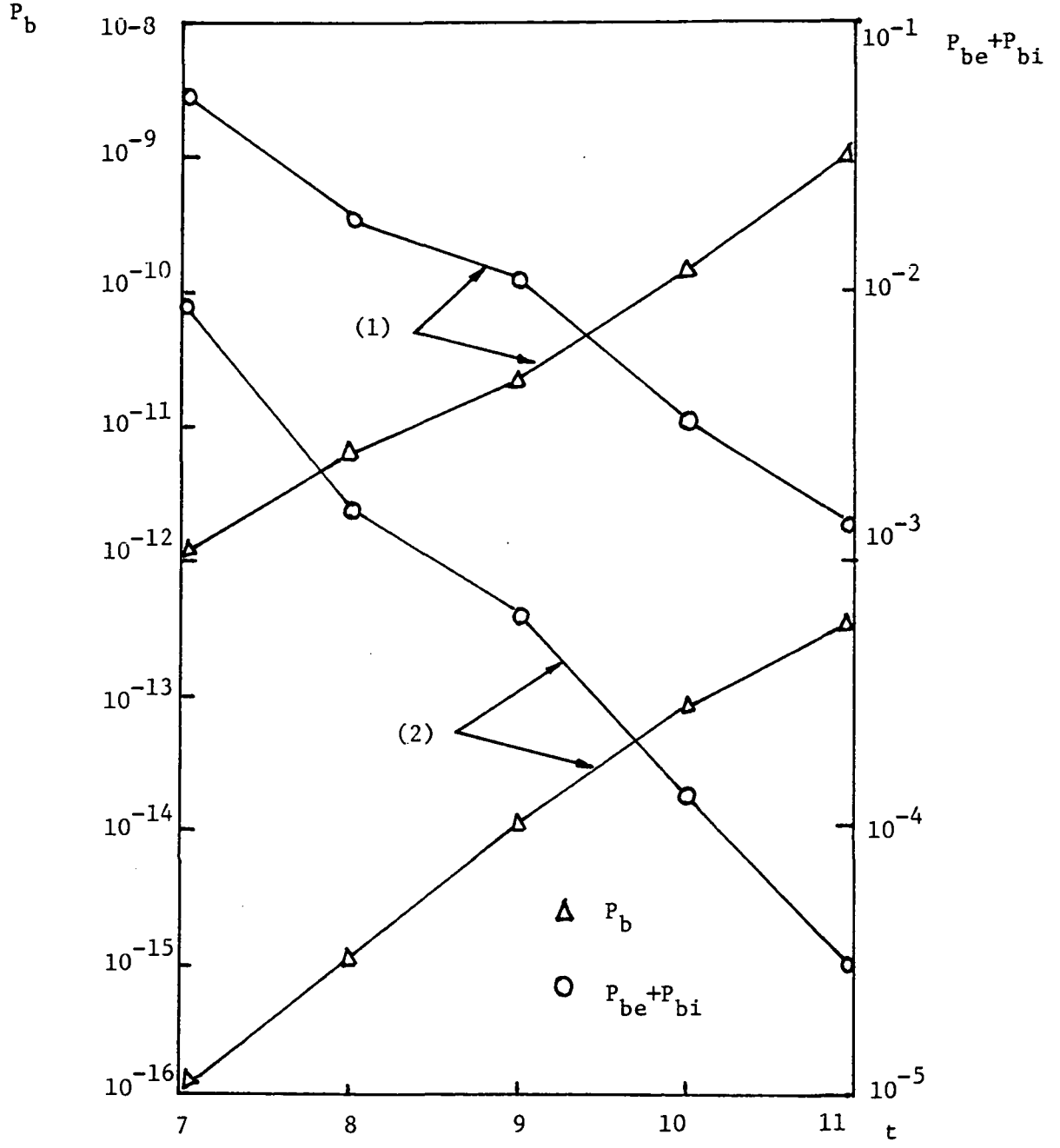

Fig. 14.1 Code performance of Example 5.1 (simulation) with R_1 = 3/4 punctured inner convolutional code, and B=4, T'=0.1, T_{es}=5, d_2=33, R=0.656. (1) E_b/N_0=3.68 dB (P_b=7.9x10^{-7} in scheme 1); (2) E_b/N_0=3.88 dB (P_b = 2.7x10^{-9} in scheme 1).

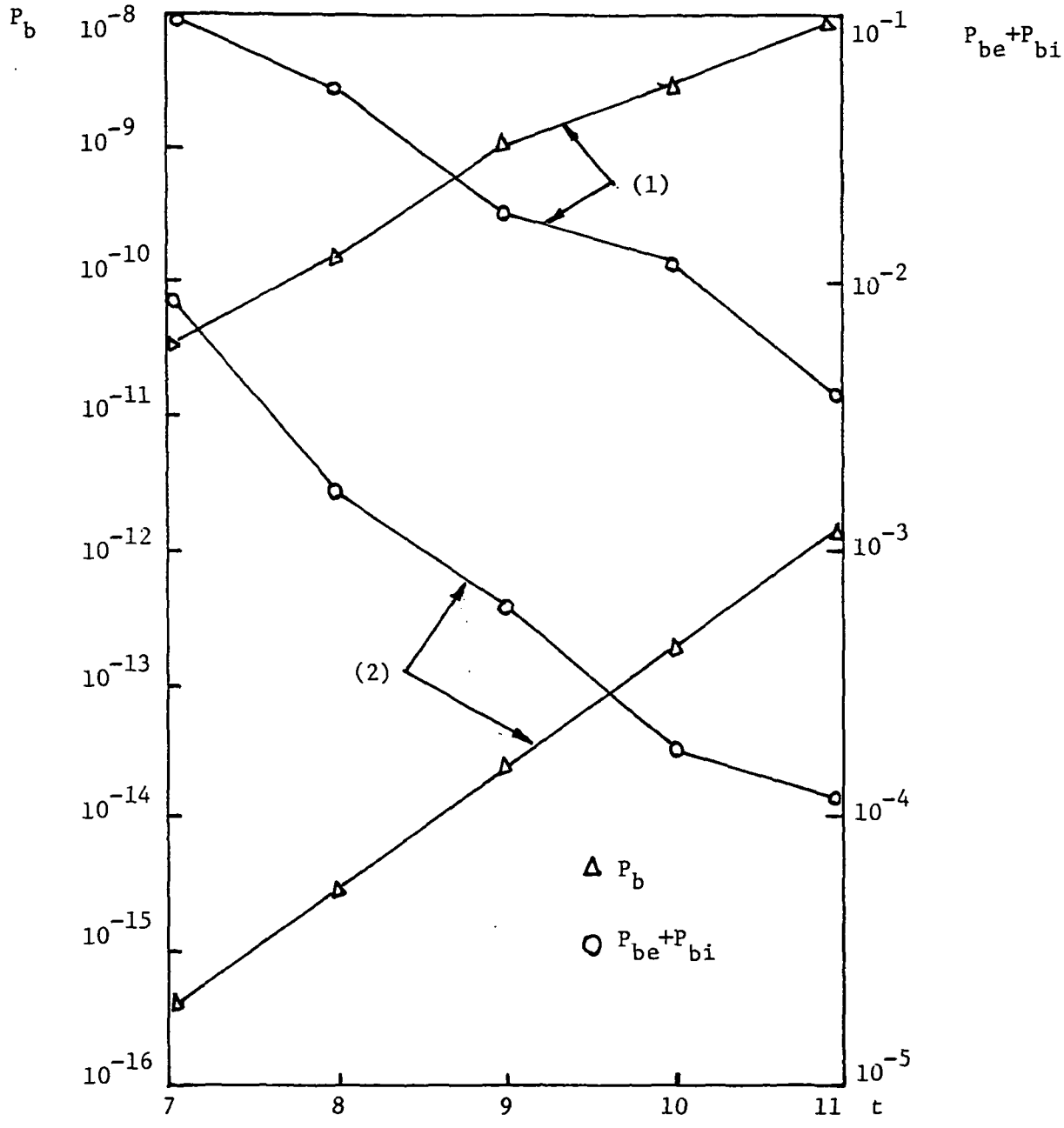

Fig. 14.2 Code performance of Example 5.1 (simulation) with $R_1=5/6$ punctured inner convolutional code and B=4, $T'=0.2$, $T_{es}=5$, $d_2=33$, R=0.729. (1) $E_b/N_0=4.56$ dB ($P_b=6.5\times10^{-6}$ in scheme 1); (2) $E_b/N_0=4.73$ dB ($P_b=4.2\times10^{-9}$ in scheme 1).

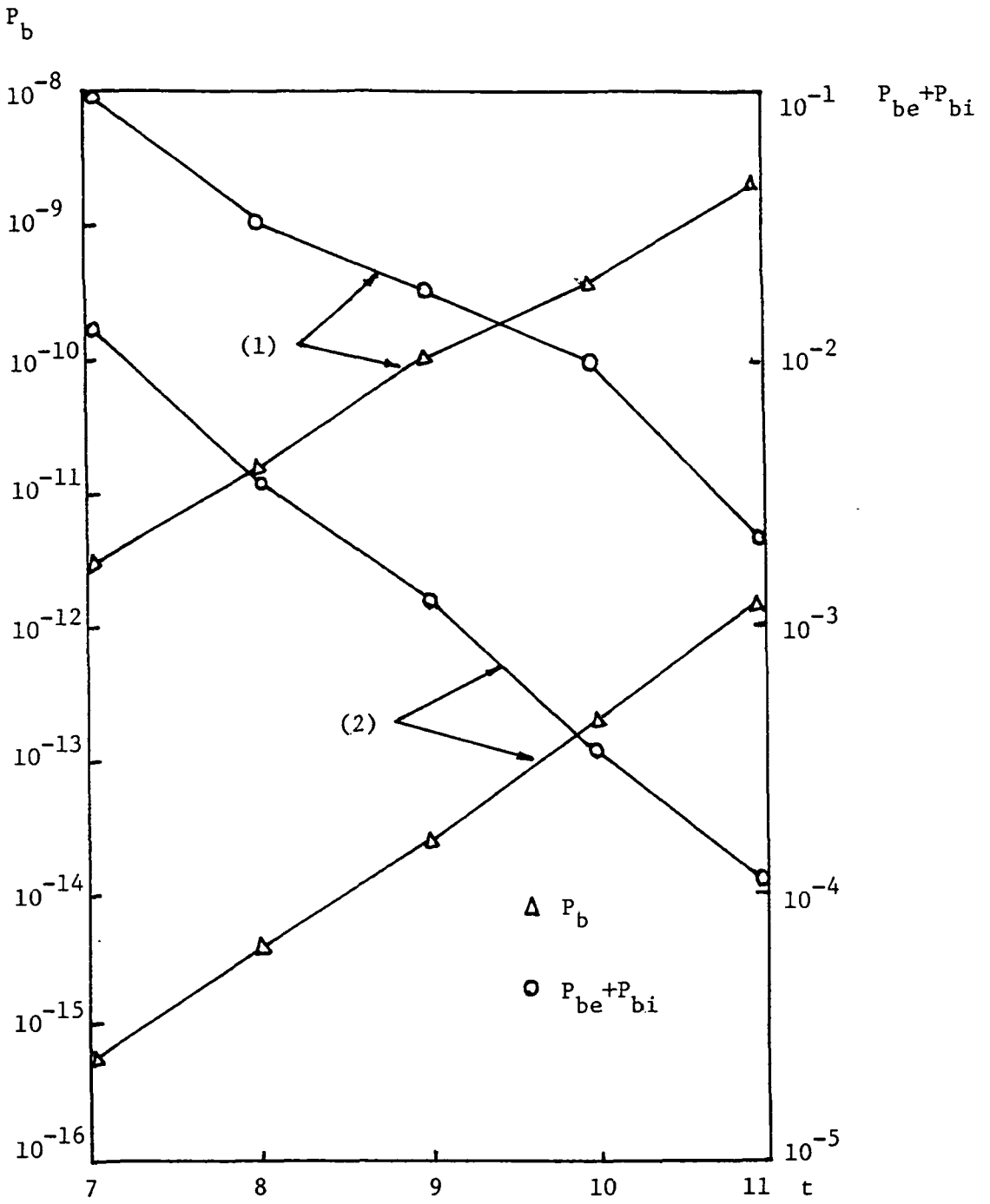

Fig. 14.3 Code performance of Example 5.1 (simulation) with $R_1=7/8$ punctured inner convolutional code and B=4, $T'=0.2$, $T_{es}=6$, $d_2=33$. R=0.765. (1) $E_b/N_0=5.20$ dB ($P_b=2.2\times10^{-6}$ in scheme 1); (2) $E_b/N_0=5.35$ dB ($P_b = 1.3\times10^{-8}$ in scheme 1).

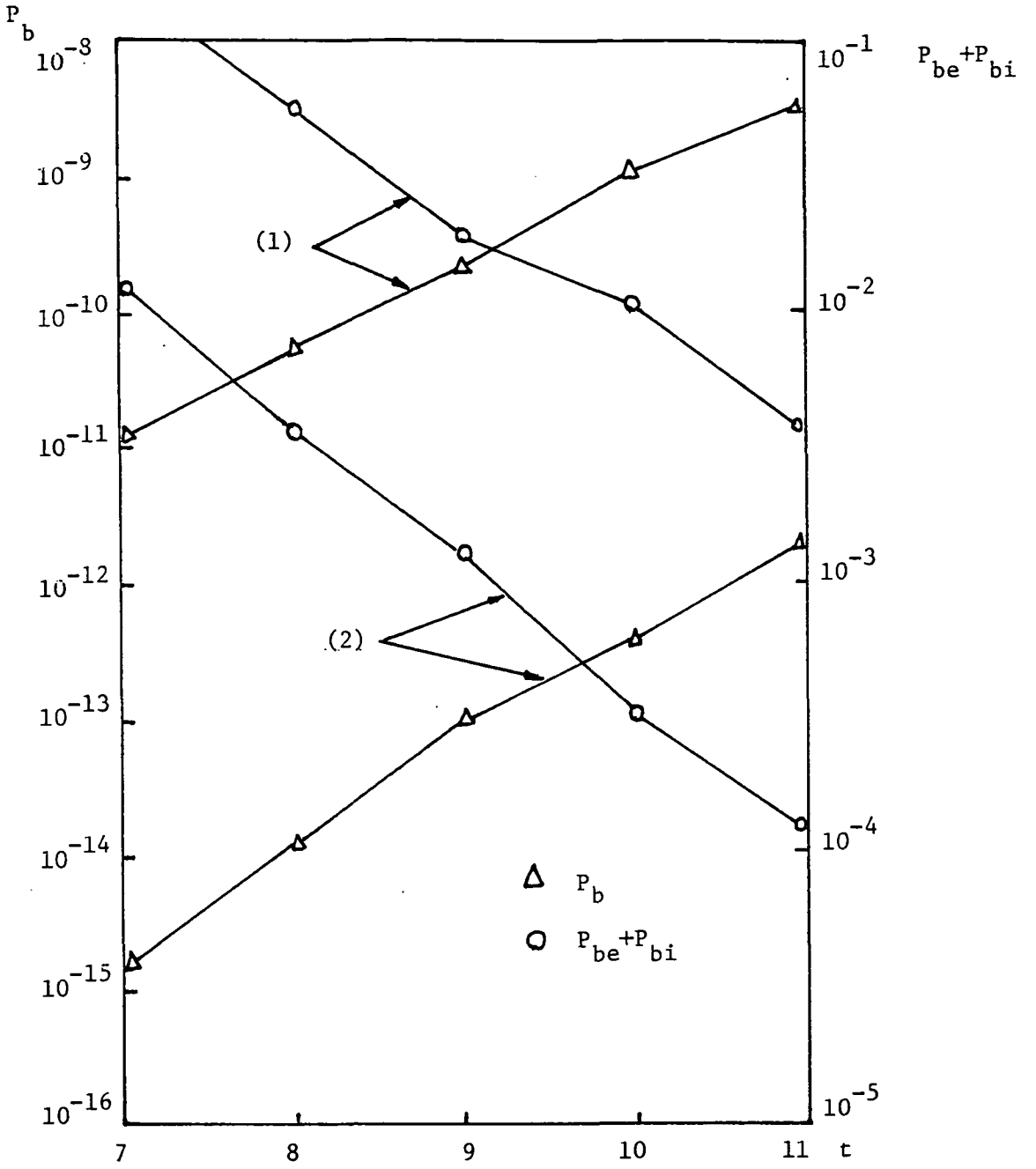

Fig. 15.1 Code performance of Example 5.2 (simulation) with 16-state R_1=7/9 PTVTC 8PSK as inner code and B=7, T'=0.2, T_{es}=10, d_2=37, R_{eff}=1. (1) E_b/N_0=6.1 dB (P_b=2.5x10^{-6} in scheme 1); (2) E_b/N_0=6.34 dB (P_b = 8.5x10^{-9} in scheme 1).

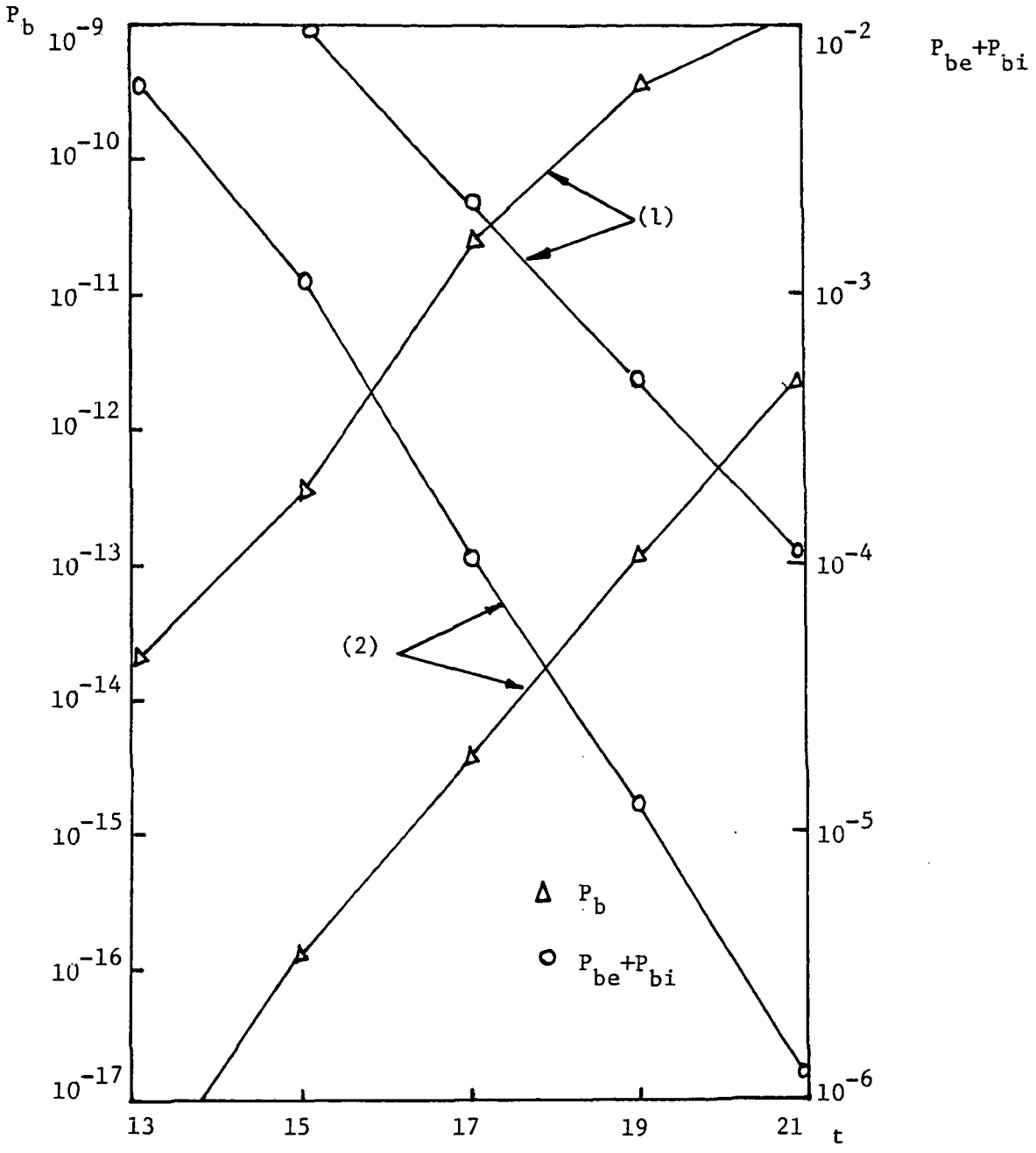

Fig. 15.2 Code performance of Example 5.2 (simulation) with 16-state R_1=5/6 PTVTC 8PSK as inner code and B=5, T'=0.1, T_{es}=8, d_2=52, R_{eff}=1. (1) E_b/N_0=6.53 dB (P_b=4.2x10^{-7} in scheme 1); (2) E_b/N_0=6.77 dB (P_n=1.9x10^{-9} in scheme 1).

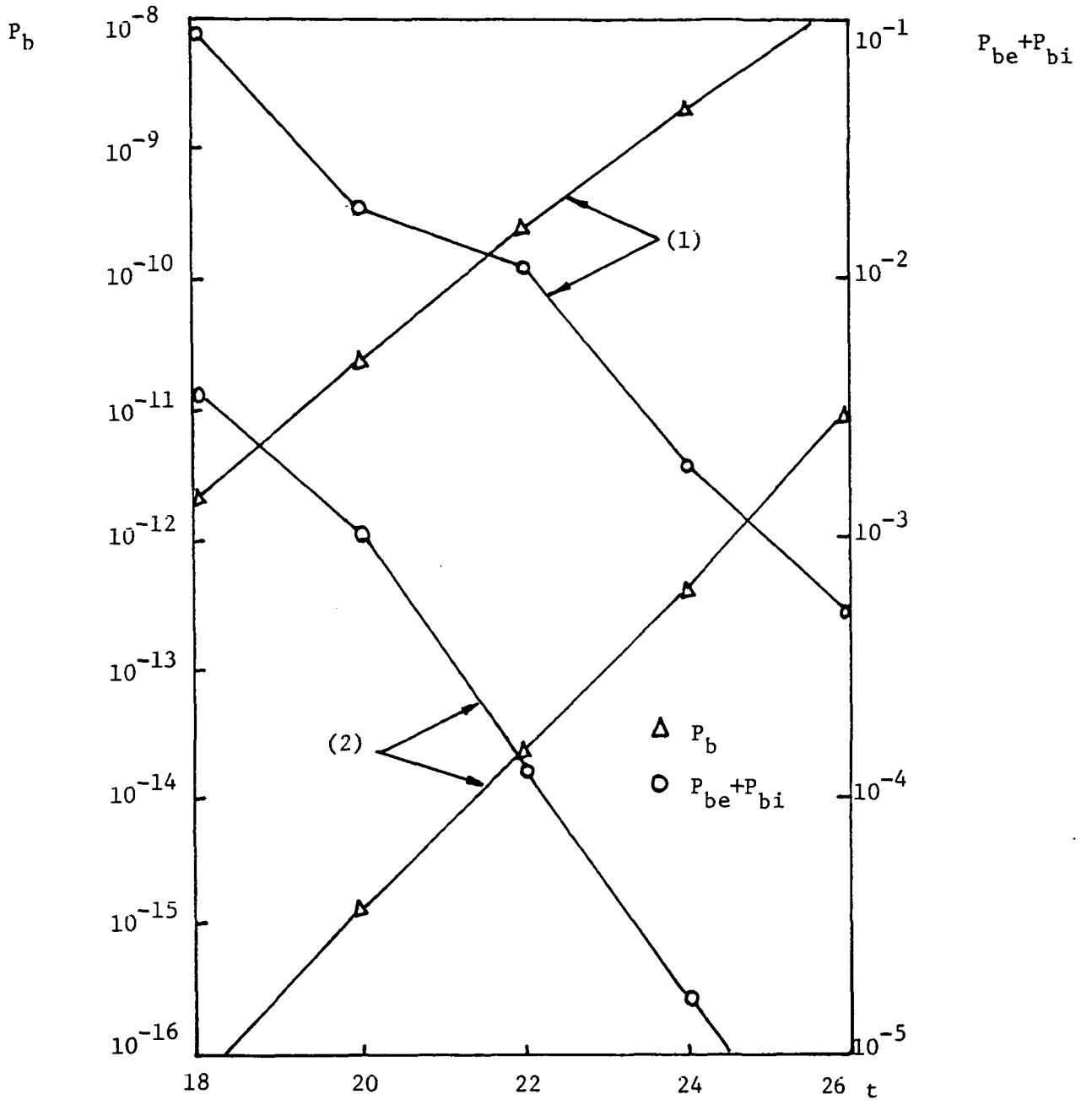

Fig. 15.3 Code performance of Example 5.2 (simulation) with 16-state R_1=8/9 PTVTC 8PSK as inner code and B=8, T'=0.1, T_{es}=9, d_2=64, R_{eff}=1. (1) E_b/N_0=7.48 dB ($P_b=4.3\times10^{-6}$ in scheme 1); (2) E_b/N_0=7.68 dB ($P_b=5.8\times10^{-9}$ in scheme 1).

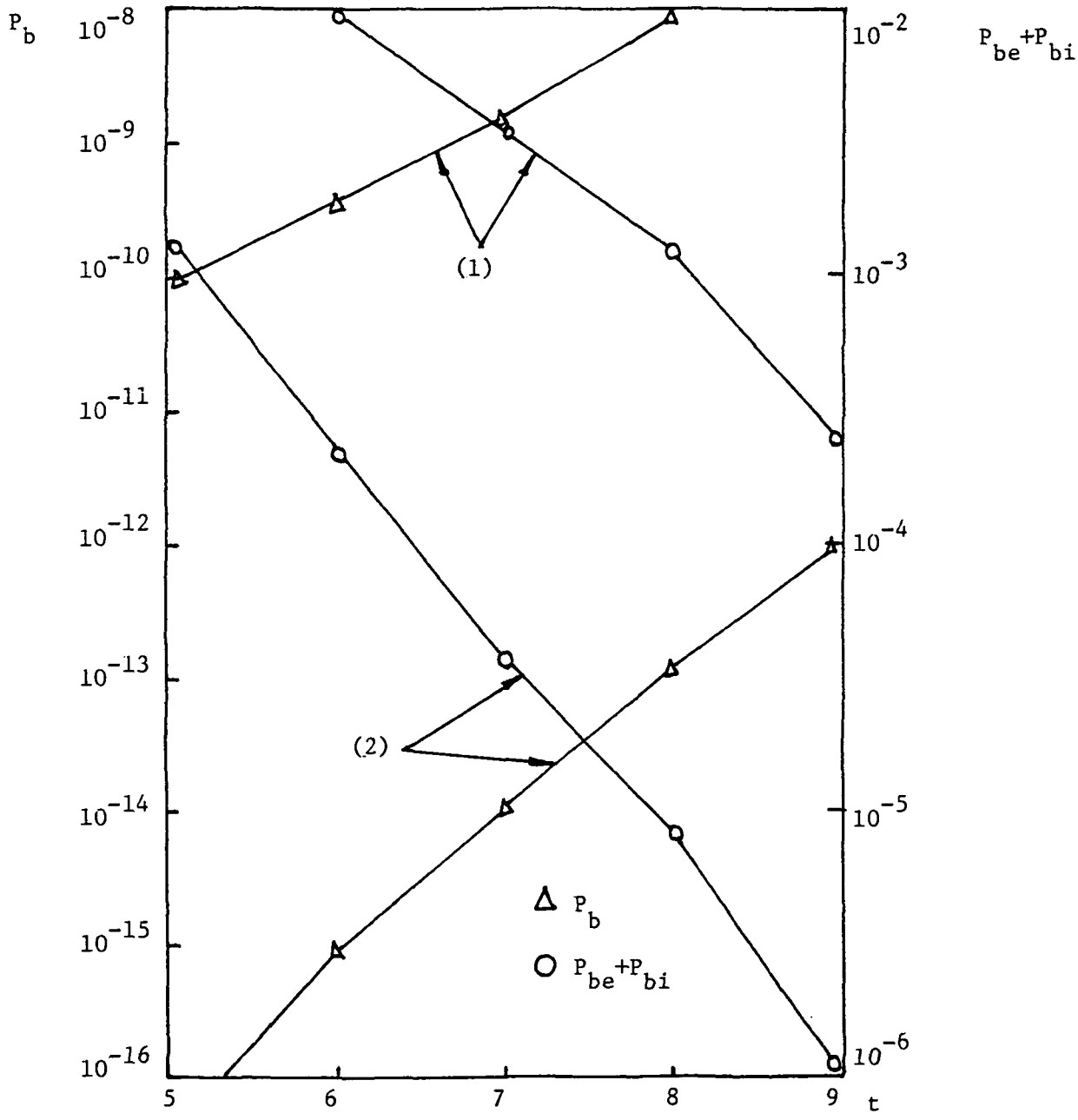

Fig. 16 Code performance of Example 5.3 (simulation) with 16-state $R_1=2/3$ coded 8PSK inner code, and B=4, $T'=0.3$, $T_{es}=10$, $d_2=26$, $R_{eff}=0.9$. (1) $E_b/N_0=5.07$ dB ($P_b=1.11\times10^{-6}$ in scheme 1); (2) $E_b/N_0=5.45$ dB ($P_b=5.3\times10^{-10}$ in scheme 1).